DOCUMENT RESUME

ED 047 927 24 RE 003 490

AUTHOR	Simons, Herbert D.
TITLE	Linguistic Skills and Reading Comprehension.
INSTITUTION	Harvard Univ., Cambridge, Mass.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-9-A-078
PUB DATE	Aug 70
GRANT	OEG-1-70-004(509)
NOTE	135p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	Cloze Procedure, *Deep Structure, Elementary School Students, Grade 5, *Language Research, Language Tests, *Reading Comprehension, *Reading Research, Structural Analysis, *Structural Linguistics, Word Recognition

ABSTRACT

The relevance of a specific linguistic concept--deep structure--to reading comprehension was investigated. The sample consisted of 87 fifth-grade students who were of above-average intelligence. To study the relationship of children's skill at recovering the deep structure of sentences to reading comprehension, the Deep Structure Recovery Test (DSRT) was designed and administered. Reading comprehension, measured by a cloze test and the reading subtest of the Metropolitan Achievement Test, correlated significantly with the DSRT. The DSRT was found to be the most important factor in reading comprehension as measured by the cloze test when compared to IQ, word knowledge, and word recognition skills. The relationship of children's skill at making a lexical analysis of the main verb of sentences and reading comprehension was measured by a sentence completion test and the subjects' performance on sets of items which contained different types of verbs. It was concluded that while the ability to recover the deep structure is an important aspect of reading comprehension, the results did not indicate a clear relationship between lexical analysis strategy and reading comprehension. Tables and a bibliography are included. (Author/DH)

BR-9-A-078
PA 24
RE

ED047927

Final Report
Project No.9A078
Grant No. OEG-1-70-004 (509)

Linguistic Skills and Reading Comprehension

Herbert D. Simons
Harvard University
August, 1970

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

RE 003 490

Final Report

Project No. 9A078

Grant No. OEG-1-70-004 (509)

Linguistic Skills and Reading Comprehension

Herbert D. Simons

Harvard University

August, 1970

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S.DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the following people. Professor Jeanne Chall, my advisor, gave me guidance and encouragement on this thesis and throughout my stay at Harvard. Dr. Helen Popp and Dr. Arthur Blumenthal provided crucial advice, criticism and encouragement on most aspects of this study. Dr. Marshall Smith with endless patience advised me on the research design and statistical analysis for this study. Dr. Thomas Hutchinson also helped with advice on statistics. Dr. Carol Chomsky aided me with suggestions for the linguistics test.

Many people in the Brookline Massachusetts School system cooperated in providing me with students to use as subjects for the study. They are Dr. Joseph Tremont, Supervisor of Language; Mr. Gerard Cote, Principal of the Lincoln School; Mr. Kenneth Allread, Miss Margot Wygant and Mrs. Florence Regolino, teachers at the Lincoln School; Mrs. Babette Raphael, Principal of the Runkel School; Mrs. Evelyn Lerman and Mrs. Cora Roth, teachers at the Runkel School.

Dr. Richard McCann of the United States Office of Education assisted me in obtaining financial support for this study.

I wish to thank Dr. Douglas Porter who was singlehandedly responsible for my coming to Harvard.

Judy Haggerty deciphered my handwriting in order to type this manuscript and this was no mean feat.

Finally I wish to thank my wife Liz whose love and faith in me plus her editorial skills made this thesis possible and my apologies

to my children, Rachel and Daniel, for all the weekends I didn't spend with them.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS.. ii

LIST OF TABLES.. vi

ABSTRACT.. viii

INTRODUCTION.. 1

CHAPTER I: RESEARCH IN READING COMPREHENSION: THE NEED FOR NEW PERSPECTIVES................................ 3

A Critique of Seven Major Approaches to Comprehension
    The Skills Approach.............................. 3
    The Measurement Approach......................... 11
    The Factor Analytic Approach..................... 14
    The Correlational Approach....................... 16
    The Readability Approach......................... 17
    The Introspective Approach....................... 19
    The Models Approach.............................. 20
    Conclusion: Approaches to Comprehension.......... 23
Comprehension Research and Pedagogy.................. 23
Why the Comprehension Process has been Resistant to Explanation...................................... 26
The Need to Base Comprehension Research on Theory.. 28
Two Types of Theories: Competence and Performance. 29
Deep Structure as a Description of the Object of the Comprehension Process......................... 32
The Psychological Reality of Deep Structure........ 34

CHAPTER II: AN INVESTIGATION OF DEEP STRUCTURE AND READING COMPREHENSION.............................. 38

The Purpose of the Study............................. 38
Hypotheses and Assumptions of the Study.............. 38

CHAPTER III: THE MEASUREMENT OF VARIABLES DESCRIBED.......... 42

The Measurement of the Variables for Hypothesis 1.. 42
    Independent Variable: Skill at recovering the deep structure of sentences.............. 42
    Dependent Variable: Reading Comprehension Skill.. 57

Page

Control Variables.............................. 62
Measurement of the Variables for Hypothesis 2...... 65
Independent Variable: Skill at making a lexical analysis of the main verb of sentences. 65
Dependent Variable: Reading Comprehension Skill.. 68

CHAPTER IV: DESCRIPTION OF THE SAMPLE AND TESTING PROCEDURES... 69

Description of the Sample.......................... 69
Testing Procedures................................. 71

CHAPTER V: RESULTS AND ANALYSIS.............................. 73

Hypothesis 1....................................... 73
Zero Order Correlations.......................... 73
Partial Correlations............................... 75
Regression Analysis............................... 77
Sex Subgroup Analysis............................. 82
Hypothesis 1: Conclusions......................... 89
Hypothesis 1: Discussion.......................... 91
Hypothesis 2....................................... 94
Sentence Completion Test.......................... 94
Item Subsets from the D.S.R.T..................... 96
Hypothesis 2: Conclusions......................... 98
Hypothesis 2: Discussion.......................... 99

CHAPTER VI: IMPLICATIONS.................................... 101

Research Implications.............................. 101
Instructional Implications......................... 102

APPENDICES

A Deep Structure Recovery Test...................... 105
B Item Statistics: D.S.R.T.......................... 111
C Cloze Tests of Comprehension...................... 112
D Regression Analyses for each School............... 115

REFERENCES... 117

VITA... 123

LIST OF TABLES

Table		Page
1	Items with Superficial Clues	50
2	Correlations between Sup. Clues and No Sup. Clues Items and Reading Comprehension	51
3	Basic Statistics of the Total D.S.R.T	56
4	Cloze Passage Calibration	59
5	Basic Statistics on Three Cloze Tests	60
6	Intercorrelations: Cloze Tests	61
7.	Basic Statistics on M.A.T. Reading	61
8	Intercorrelations: Dependent and Control Variables	64
9	Subset of D.S.R.T. Items: Transitive Verbs-Complement Verbs	67
10	Subset of D.S.R.T. Items: Tell vs. Ask	68
11	Basic Statistics of the Sample	69
12	Basic Statistics of the Sample by Sex	70
13	Correlations between D.S.R.T. and Reading Comprehension	73
14	Intercorrelations of all Variables Used in Testing Hypothesis 1	76
15	Partial Correlations: D.S.R.T. and Reading Comprehension with I.Q., Word Discrimination (W.D.), Word Knowledge (W.K.), controlled	76
16	Regression Analysis: D.S.R.T., Word Knowledge, Word Discrimination, I.Q. with Cloze Test	79
17	Regression Analysis: D.S.R.T., Word Discrimination, Word Knowledge, I.Q. with M.A.T. Reading	80

Table		Page
18	Comparison of Means for Males and Females	82
19	Correlations between D.S.R.T. and Comprehension by Sex	83
20	Partial Correlations: D.S.R.T. and Reading Comprehension with I.Q., Word Discrimination (W.D.), Word Knowledge (W.K.) controlled. Analysis by sex	84
21	Intercorrelations of All Variables used in Testing Hypothesis 1, by Sex	86
22	Regression Analysis by Sex: D.S.R.T., Word Knowledge Word Discrimination, I.Q. with Cloze	87
23	Regression Analysis by Sex: D.S.R.T., Word Knowledge Word Discrimination, I.Q., with M.A.T. Reading	88
24	Correlations: Sentence Completion Test vs. Reading Comprehension	95
25	Correlations: Sentence Completion vs. Reading Comprehension by Sex	95
26	t-tests, Simple vs. Complex Verbs by Reading Comprehension (Upper 40%/Lower 40%)	97

ABSTRACT

This study is concerned with the relevance of a specific linguistic concept -- deep structure -- to reading comprehension. In linguistic theory it is the deep structure which describes the logical subject and object of a sentence. The relevance of this concept (deep structure) to reading comprehension is first studied by looking at the relationship between children's skill at recovering the deep structure of sentences and their reading comprehension. Then, one strategy that children might use in recovering deep structure and the relationship of this strategy to reading comprehension is studied. This strategy -- the lexical analysis strategy -- involves analysis of the main verb of a sentence to determine the deep structures with which it is compatible.

The sample for the study consisted of 87 fifth grade students from a suburban community. The students were of above average intelligence with a mean I.Q. of 117.

The first hypothesis of the study is: Children's skill at recovering the deep structure of sentences is related to reading comprehension skill. This hypothesis is evaluated through a test, the Deep Structure Recovery Test (D.S.R.T.), which is designed to measure Ss' skill at recovering the deep structure of sentences. Each item of the D.S.R.T. contains three sentences. In taking the test an S is asked to choose the one sentence out of the three which is not a paraphrase of the other two sentences. In each item two of the sentences

have the same deep structure while a third sentence, although superficially similar to at least one of the other two sentences, has a different deep structure and thus a different meaning, for example,

*(a) What the boy would like is for the girl to leave.

(b) For the boy to leave is what the girl would like.

(c) What the girl would like is for the boy to leave.

In this study reading comprehension is measured by a Cloze test. It is argued that a Cloze test is superior to traditional standardized comprehension tests. The traditional Metropolitan Achievement: Reading Subtest (the M.A.T. Reading test) is also used for comparison purposes. I.Q., word knowledge and word recognition ability are also measured. A correlation and regression analysis produced these results:

1. The D.S.R.T. has a substantial and significant correlation, i.e. .732, with the Cloze test and a significant correlation with the M.A.T. Reading test.
2. The D.S.R.T. is the most important factor in reading comprehension as measured by the Cloze test, when compared to I.Q., word knowledge and word recognition skill. The D.S.R.T. alone accounted for more than 20% of the unique variance in reading comprehension. The other variables made little if any unique contribution.

---

*The asterisk indicates the sentence with the different deep structure.

3. Word knowledge is the most important factor in reading comprehension as measured by the M.A.T. Reading test. However neither word knowledge nor any of the other variables contributed much unique variance.

The second hypothesis of this study is: Children's skill at making a lexical analysis of the main verb of sentences is positively related to reading comprehension skill. Ss' skill is measured in two ways, first by a sentence completion test and second by Ss' performance on sets of items (included in the D.S.R.T.) which contained different types of verbs. Reading comprehension again is measured by a Cloze test and the M.A.T. Reading test. While the results of the analysis of the Sentence Completion test support the hypothesis, the results of the analysis of the items on the D.S.R.T. do not.

In short, it appears that Ss' ability to recover the deep structure of sentences is an important aspect of reading comprehension. It is not clear, however, that the lexical analysis strategy is related to reading comprehension.

And so to completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind, as well as to unravel the tangled story of the most remarkable specific performance that civilization has learned in all its history.

Edmund Burke Huey

THE PSYCHOLOGY AND PEDAGOGY OF READING

INTRODUCTION

As Thorndike pointed out in 1917 reading comprehension is "...a very complex procedure involving a weighing of each of many elements in a sentence, their organization in proper relation to one another, the selection of certain elements of their connotations and the rejection of others and the cooperation of many forces to produce the final response [1917, p.323]." Unfortunately, in spite of the vast literature produced on the topic of reading comprehension over the past 50 years, this quote almost completely exhausts the accumulated knowledge of this fundamental intellectual process. Despite this voluminous research the workings of the mind during reading comprehension remain a great and profound mystery. Consequently, the instructional procedures and materials used when teaching children to comprehend what they read are based upon the intuitions and accumulated experience of reading specialists, not on research evidence.

The major reason, beyond methodological problems, for the inadequacy of past research has been a lack of basic knowledge of the psychological processes involved in reading comprehension. If instruction in reading comprehension is to be based on research evidence as well as the intuitions and experience of reading specialists, it is important that fruitful research shed some light upon the psychological processes involved in reading. Knowledge of these processes can then provide the basis for instruction.

It is the purpose of this thesis to suggest a new approach to the study of the reading comprehension process and to report an empirical investigation of the author's which has been motivated by this new approach. Fundamental to this new approach is the need to base comprehension research on theory. The linguistic theory developed by Chomsky (1957, 1965) is the one on which this thesis is based. The study reported in this thesis is an investigation of the relationship of one aspect of this linguistic theory to the reading comprehension process.

The paper is organized into six chapters. Chapter One offers a critique of past research and a discussion of the direction that research, in the author's opinion, must take if it is to illuminate the processes basic to reading comprehension. Also in this chapter the rationale for basing this study on linguistic theory is developed. The hypotheses and assumptions of the study are presented in Chapter Two. Chapter Three is a description of the measurement of the variables used in the study. The sample studied and the testing procedures are described in Chapter Four. The results and the analysis of the data are contained in Chapter Five. And finally in Chapter Six the implications of the study are discussed.

CHAPTER I

Research in Reading Comprehension: The Need for New Perspectives

A Critique of Seven Major Approaches to Comprehension

In this chapter the inadequacies of the past research will be discussed, and the direction that research should take in order to lead to an understanding of the complex process of reading comprehension will be described.

The Skills Approach

Many reading specialists believe that reading comprehension can best be described and understood through the enumeration of lists of comprehension skills. Therefore, most of the research has been influenced directly or indirectly by this belief in a skills explanation. Much of the research has endeavored to measure and verify existing skills and discover new ones. A discussion of these skills and the problems they raise will help demonstrate why the research has been so unproductive.

A great many lists of skills have been proposed, with some lists containing several hundred or more skills. The following is a sampling of skills that have been proposed:

- reading in thought units
- grasping and assimilating relevant details
- fusing new and old ideas
- thinking about words

- seeing relationships between words and thoughts
- remembering what is read
- finding the topic sentence
- making generalizations
- drawing inferences
- underlining the key words in a paragraph
- predicting outcomes
- distinguishing fact from opinion
- recognizing literary devices
- evaluating the author's competence
- combining recall with own associations
- reading to follow directions
- offering new titles for paragraphs
- finding the main idea of a paragraph

Recently some of the lists of skills have been organized into taxonomies (Spache, 1962; Cleland, 1965; Wolf et al, 1968; Barrett, 1968). Taxonomies are presumed to be superior to simple lists because they provide systematic organization which is intended to increase our understanding of the phenomenon under investigation. Taxonomies are systematic in the sense that they organize skills into major categories and subcategories. For example in the Barrett (1968) taxonomy one major category is inferential comprehension and the subcategories are inferring supporting details, inferring main ideas, inferring sequences, inferring comparisons, inferring cause and effect relationships, inferring character traits, predicting outcomes and interpreting figurative language.

The advantage of systematic organization, however, is more imagined than real. As Clymer (1968) has pointed out "as is true with all taxonomies, the orderly presentation of the categories may suggest a greater precision than the classification system really possesses [p.19]." Clymer's observation applies to taxonomies of reading skills. The analysis that follows applies equally to taxonomies and simple lists of skills.

The skills approach to an explanation of reading comprehension suffers from a basic confusion over what domains of behavior and cognitive activity can rightfully be characterized as reading comprehension. This confusion has led to the inclusion of reading comprehension skills that tend to be global and vague and that fail to distinguish between those processes that are specific to reading and those that are very general cognitive processes.

The lists and taxonomies of comprehension skills fail to clearly distinguish between general mental abilities and those skills specific to reading. For example, the reading skill of "making generalizations" is a skill common to most cognitive processes and hardly specific to reading. In designating such of the thinking skills as reading comprehension skills the reading specialists are now obliged to explain the thinking process in order to include the thinking skills in reading comprehension instruction. To say that this is a formidable task is an understatement since very little is presently known about thinking processes.

Another problem with lists and taxonomies of comprehension skills is that they fail to make the distinction between how something is comprehended and what is comprehended. As a result they fail to define adequately either how something is comprehended or what is comprehended. For example, in the skill of "fusing new and old ideas" what is learned is "the ideas" and the process is the fusion of these ideas. In order for the description of this skill to be meaningful, it is necessary to provide an adequate theory of what ideas are as well as a specific description of the fusion process. Neither the adequate theory of ideas or a specific description of the fusion process are presently available. This case is typical of the status of many comprehension skills. It appears that a description of what is comprehended is necessary before attempting to describe the process of how something is comprehended. So that in the skill of "fusing new and old ideas" a theory of ideas must be developed before research into the "fusion" process can begin. It looks like the place to begin in comprehension research is with theories of what is comprehended. The recent work in linguistic theory (Chomsky, 1957, 1965) offers just such an opportunity. This point which has motivated the present study will be expanded upon later.

Another problem with mental skills as explanations of reading comprehension is that they have been used in the literature to describe three distinguishably different types of activities as follows:

1. High level cognitive processes, i.e., processes that involve very sophisticated thinking abilities which go beyond a literal understanding of the reading matter. These high level cognitive skills include: "distinguishing fact from opinion", "recognizing literary devices", "evaluating the author's competence", etc.

2. The psychological processes involved in reading comprehension, "fusing new and old ideas", "reading in thought units", "combining recall with own associations."

3. The procedures for teaching comprehension e.g., "finding the topic sentence", "offering new title for paragraphs", etc.

Unfortunately, reading specialists have often failed to distinguish between these three categories in their discussion of reading comprehension skills. Consequently, there has been a great deal of confusion[1] as reading specialists attempt to evaluate the importance of the various skills and develop ways of teaching these skills to children. This is the case because in attempting to evaluate the importance of the proposed skills and in developing techniques for teaching these skills, there are different considerations to be taken into account depending upon which of the above category of skills one is talking about. This point will become clearer as each category of skill is discussed in turn.

---

[1] Some skills fall into more than one category depending upon how they are interpreted. For example, the skill "thinking about words" could be either a high level cognitive skill or a process skill. The fact that some skills have more than one meaning, of course only adds to the confusion.

The first category is high level cognitive processes. In evaluating the importance of these skills the most important consideration is whether enough is understood about the mature reader's skill to teach them to children. This will involve an understanding as discussed above of both the objects of the skill as well as the process employed in using the skill. Very few high level cognitive skills have been analyzed in light of this consideration. However, Schell (1967) has analyzed the high level cognitive skill of "distinguishing fact from opinion" along these lines. He shows first that the task of actually distinguishing fact from opinion is a very difficult task much of the time, because the distinction between fact and opinion is so ill-defined that in all but the most obvious cases this distinction is hard to draw.

In addition, in the cases where this distinction could be drawn the relevant information for drawing this distinction is often not contained in the reading matter. Thus, in the case of the comprehension skill of "distinguishing fact from opinion" it appears that we do not know very much about the objects of this skill, i.e., facts and opinions, nor about the processes employed in applying this skill. Consequently it is not surprising that Schell (1967) finds that "there are no handy rules of thumb we can give students to help them in this task [p.9]." The rule of thumb now used for teaching children how to decide when something is an opinion is to look for qualifiers, e.g., seems, appears, may, probably, or indicators, e.g., I think, we believe, our conclusion, etc. This rule of thumb is of questionable validity because authors commonly omit qualifiers and indicators in expressing opinions. Thus the lack of knowledge about the object and the processes of comprehension reduces the teaching of this skill to superficial techniques of questionable validity. Examination of other higher level cognitive skills reveals more of the same.

The second category of skill presumes to describe the psychological processes employed in comprehension. Since, as will be argued in this chapter, description and explanation of the psychological processes involved in reading comprehension are a crucial prerequisite to improving instruction, the validity of these process skills is of utmost importance. The question to be answered in assessing the validity of process skills is, how good a description and explanation do these skills provide of the comprehension process? In examining the process skills one is forced to conclude that the answer to this question is that the process skills proposed in the literature do not describe and explain the comprehension process at all. Such skills as "reading in thought units," "fusing new and old ideas," "grasping and assimilating relevant details" hardly qualify as precise descriptions of the comprehension process. They are merely metaphors and as such do not lend themselves to scientific theory construction or empirical research. These skills, or rather the present descriptions of them, only raise a lot of questions without providing any answers. Wardhaugh (1969) in talking about the fusion metaphor gives a succinct statement of some of the problems raised by this type of vague description.

> ...the fusion metaphor itself is a poor and unrevealing one in that it short circuits the process of explaining exactly what happens when someone understands a sentence. The concern should be with exactly what factors are involved in fusion and why fusion occurs on some occasions but not on others [p.86].

For the process skills to be meaningful as descriptions of the comprehension process it is necessary first to know more about the objects of these skills, i.e., ideas, thoughts, details, etc. In addition, it is necessary to provide an explicit description of the fusion, assimilation, and grasping process, etc., and the conditions under which they are utilized. Only then can skills such as "fusing new and old ideas" have any significance as descriptions of the comprehension process.

In many lists and taxonomies one finds comprehension skills such as "underlining the key words in a paragraph," "finding the topic sentence," "offering new titles for paragraphs" etc. This group of "skills" has caused some confusion because they are quite clearly teaching procedures to be used to improve comprehension rather than descriptions of the comprehension process. Sometimes, however, these skills are used as testing procedures to demonstrate that comprehension has taken place. In any case they are really a separate category and should not be listed as comprehension skills at all.

It appears that simply because they were listed along with comprehension skills, they are often treated as ends in themselves.

Thus teachers dogmatically insist that students must learn to underline the key words in a paragraph etc., if they can't learn to do this they are considered to have poor comprehension skills. But since the "skills" are means rather than ends their validity and importance depends upon how successful these means (i.e., teaching procedures) are for achieving the desired ends (comprehension). This of course assumes that comprehension is well understood. As argued above, not very much is known about comprehension skills so that there is no real way of evaluating these teaching procedures, and they must be considered only tentative at the present time.

To summarize, in the discussion up to this point it has been argued that the setting up of categories of skills approach has not shed much light upon reading comprehension because of a basic confusion over the precise behavior and cognitive domain of these skills. This confusion has led to global and vague skills which have failed to distinguish (a) between reading and thinking, (b) between the objects and the processes of comprehension, and (c) between the high level cognitive processes, psychological processes and teaching procedures of comprehension.

The Measurement Approach

The most popular way of measuring comprehension has been through the development of standardized reading comprehension tests. Typically these tests have been designed to measure successful performance on the comprehension skills discussed in the previous section. In the most common format the student is asked to read a passage and then answer a series of questions which require him to select from alternative choices the main idea of the passage, or find some details, or select the predicted outcome, or select a generalization, etc.

There is a fundamental problem with traditional reading comprehension tests. Simply stated, it is not at all clear what they are measuring. They have been criticized because they may be measuring in addition to or in place of comprehension:

1. a student's memory skill on tests where the student is not allowed to look back at the passage (Davis, 1962).
2. a student's ability to comprehend the questions and alternative choices in the test.
3. a student's test taking skill.
4. a student's familiarity with the content of the passage. Kerfoot (1965) has shown that many of the questions on comprehension tests can be answered without reading the passage.
5. a student's motivation and attitude.
6. a student's personality.
7. a student's ability to recognize the words.

The lack of a theory of the reading comprehension process makes it almost impossible to separate the above aspects of a student's functioning from his reading comprehension skill. For this reason traditional reading comprehension tests lack construct validity (Chronbach & Meehl, 1955); that is, no clear relationship has been established between the test and the underlying concept -- reading comprehension. Until theories of reading comprehension are developed, tests with construct validity will be impossible.

In recent years a newer and better method of measuring reading comprehension has been developed. It is the Cloze procedure (Taylor, 1953). On a Cloze test every nth word of a passage is deleted and a student takes the test by filling in the blanks where the words were deleted. The Cloze test is a better measure of reading comprehension than traditional tests because it is measuring reading comprehension, i.e. it correlates with traditional comprehension tests and at the same time it appears to be measuring fewer of the extraneous aspects of student functioning. Specifically, it does not have questions and therefore is not measuring a student's skill in understanding questions. It is not a memory test because the student can continually re-examine the passage. It also does not appear to be measuring a student's familiarity with the content of the passage, at least to the degree that traditional tests do. The mechanical procedure for developing the test (the deletion of every nth word) renders it more objective and less subject to the arbitrary judgments of the test constructor than traditional comprehension tests. In addition, a great deal of research has gone into the Cloze test and much is known about its operating characteristics (see Tremont, 1967; Potter, 1968). Obviously the Cloze test lacks construct validity for the same reason that all tests of the reading comprehension process do, that is, the absence of a theory of the process.

In summary, traditional and Cloze tests both suffer from a lack of construct validity. However the Cloze test is a better measure of comprehension because in measuring reading comprehension it appears to be measuring fewer extraneous aspects of cognitive functioning than traditional tests do.

The Factor Analytic Approach

Factor analysis is another technique that has been employed in an attempt to understand reading comprehension. This approach has been closely related to skills and their measurement by standardized tests because factor analysis has most often been applied to standardized tests.

There have been a number of factor analytic studies of reading comprehension (Davis, 1944, 1968; Thurstone, 1946; Vernon, 1957; Hall & Robinson, 1945; Anderson, 1949; Langsam, 1941; Gans, 1940; Holmes & Singer, 1966). Since the results of a factor analysis are in part determined by the age of the subjects, the tests used, the mathematical solution employed and the labels applied to the factors, and since these varied from study to study, it is not surprising that we find some very disparate factors such as seeing relationships, word meaning, number factors, chart reading skills etc., in these studies.

In general, these studies found that reading comprehension was composed of the same skills that reading comprehension tests measure -- not an earth-shaking discovery. In examining the factor analytic studies for the factors that showed up in a number of studies, Spache ([illegible]) found three [illegible] factors in reading comprehension; a word meaning factor, an idea relationship factor, a reasoning factor. But these factors tell us very little because they are no more than metaphorical descriptions of the reading comprehension process -- the same problem that existed before these studies.

Davis (1944) points out a requirement that factor analytic studies of comprehension must meet in order to be meaningful:

> The most important step in a study that employs factorial procedures for the investigation of reading comprehension is the selection of the tests the scores of which are to be factored. Unless these tests provide measures of the most important mental skills that have to be performed during the process of reading, the application of the most rigorous statistical procedure cannot yield meaningful or significant results. The importance of this point can hardly be overestimated [p.3].

Clearly, comprehension tests do not meet the requirement that they measure "the most important mental skills that have to be performed during the process of reading," because the exact nature of these "mental skills" is not yet known. The task for reading researchers is to discover and describe these "mental skills" in a very precise way. Only then can tests be developed to measure these skills and then factor analysis may become a useful tool in looking at the relationships between these "mental skills."

The Correlational Approach

Another widely used approach to understanding reading comprehension has been through the search for correlates of the comprehension process. A number of studies, employing standardized tests as measures of comprehension and correlational or group comparison research designs, have been conducted to determine what variables are related to the comprehension process. These studies have found such correlates of comprehension as:

1. Social class (Chandler, 1966)
2. Race (Cooper, 1964)
3. Sex (Gates, 1961)
4. Personality (Vehar, 1962; Neville, Pfost & Dobbs, 1967)
5. Attitude (Jacobson & Johnson, 1967)
6. Physical growth (Olson, 1940)
7. Intellectual ability (Harootamian, 1966; Bleismer, 1954)
8. Perceptual skill (Olson, 1966)
9. Rate of reading (Tinker, 1939; Carlson, 1949)
10. Oral language skill (Tatham, 1967; Loban, 1966)
11. Listening (Duker, 1965)

These studies and others like them raise the same problems. Most, if not all, of these studies employed standardized reading tests as measures of their dependent variable, i.e., reading comprehension, thus they ipso facto define reading comprehension in terms of the skills these tests purport to measure. However, as has been pointed out, the skills and their measurement are of questionable validity, thus the relationships found by these studies between the above variables and reading comprehension are difficult to interpret, because the dependent variable of the studies is of questionable validity.

It is difficult if not impossible to determine what the factors enumerated above correlate with, since it is not clear what comprehension tests are measuring. Thus I.Q., race, sex, etc. may be correlated not with the reading comprehension process but with the ability to answer questions, test taking skill, motivation and the like.

But more importantly, even if the relationships found in these studies prove to be correct, how much does this add to our knowledge of reading comprehension? It clearly is of some value to know that intellectual ability, language ability, perceptual ability, etc., are related to reading comprehension. But the question of how and in what ways these variables are related to comprehension remains unanswered. What do these relationships indicate about the comprehension process? When more is found out about the comprehension process, then the findings of these studies will be of more value.

In a sense these studies provide some of the facts that a theory of comprehension must explain. However enough of these facts have been gathered for the present time. As Wardhaugh (1969) puts it:

> ...these factors multiply faster than systems for handling them, so that we are very far at the moment from an acceptable theory of reading. This situation will continue to exist as long as researchers allow themselves to be sidetracked from the content of reading into its correlates [p.4].

The Readability Approach

Readability research is another area that is closely related to searching for correlates of the comprehension process. Looking at it from the point of view of trying to understand comprehension, readability research is the search for those characteristics of written matter that are correlates of reading comprehension. Chall (1958) presents an exhaustive review of the work in readability. More recent developments can be found in Klare (1963) and Bormuth (1968). This research has been productive because it has had a direct influence on instruction. Publishers and teachers have used these formulas extensively in developing and editing reading materials for children. Thus readability research has had practical consequences and this in and of itself justifies the research.

But from the point of view of understanding reading comprehension the question to ask is; how much light has readability research shed upon reading comprehension? The answer is some but not much.

Since readability research has tended to use comprehension tests that measure the same skills that have been discussed above, the characteristics of written matter that have been found in these studies to correlate with comprehension difficulty can be questioned because the measurement of the dependent variables of these studies lack construct validity.[2] Even if it is assumed that the dependent variables of these studies are accurately measured by comprehension tests, the question to be answered is; what properties of printed matter account for comprehension difficulty and how informative are these properties? Chall (1958) has suggested that there are four major factors that account for comprehension difficulty: (a) vocabulary load, (b) sentence structure, (c) idea density and (d) human interest. Idea density and human interest are so global and fuzzy that they haven't been measured very successfully. Vocabulary is measured most accurately by the number of words in a passage not on a given list of frequent words, and sentence structure is most accurately measured by some measure of sentence length. These factors appear to be approximate measures of some underlying variables that are intrinsic to the comprehension process. In order for these variables to help in understanding reading comprehension the processes underlying them must be explained. An understanding of the reading comprehension process will then answer questions such as: What is it about sentence structure and vocabulary load that influences comprehension difficulty? In what ways do sentence structure and vocabulary influence comprehension? What are idea density and human interest and how can they be described more precisely etc.?

---

[2] The more recent readability studies have used the Cloze test to measure comprehension. As has been pointed out the Cloze test is superior to the traditional measures in ways already discussed. It does, however, still lack construct validity.

The Introspective Approach

Another approach to understanding comprehension has been the study of readers' introspective reports immediately after reading (Cafone, 1966; Piekarz, 1968; Jenkinson, 1957). This type of study is admirable in that it proposes to take a close look at the comprehension process and the mental activities involved in reading.

However, introspective studies suffer an important methodological limitation in the case of reading comprehension. Since reading is a covert and complex mental process which takes place at great speed, it does not appear to be open to introspection. If one thinks about reading when one is reading, then the comprehension process is turned off, so that one can't describe the process during reading because of an inability to pay attention to two things at once. It is not at all clear what the reports of subjects in these studies are descriptions of. There are so many factors such as memory, thinking, personality, motivation, etc., other than the reading comprehension process that could account for these descriptions that the relationship between these descriptions and the comprehension process is questionable. In order for these studies to bear fruit it must first be shown that the descriptions that readers produce of their own reading bears a close relationship to their actual reading process. And given the speed, complexity and covert nature of the comprehension process the likelihood that these descriptions bear any close relationship to the reading comprehension process seems remote to this investigator.

The Models Approach

A recent approach to understanding the comprehension process has been the construction of theoretical models of the process (Spache, 1962;

Smith, 1960; Kingston, 1961; Holmes & Singer, 1966. As Clymer (1968) puts it:

> Constructing a model forces the investigator to organize facts and to set them against a rational framework; at the same time, it provides a technique for testing these facts and for generating more hypotheses for testings [p.12].

It seems clear that models are one fruitful way of understanding the comprehension process. A model of reading comprehension can generate testable hypotheses which can be verified through empirical research and can ultimately lead to an adequate theory of the process.

The question that must be asked about the proposed models of the comprehension process is; how much light do they shed upon the comprehension process? Unfortunately, the promise of the model's approach to reading comprehension has not been fulfilled. The models cited above of reading comprehension fall short because:

1. The components and categories of these models tend to be vague and global. The major terms are undefined. They look very much like the same old comprehension skills. Thus we find such model components as "recognition of a sentence as a complete thought," "comprehend main idea as extension of the topic sentence" (Spache, 1962) and "main idea as implied or as topic sentence" (Smith, 1962). The problems with these types of statements whether as skills or as components of models should be obvious.

2. There is a confusion of psychological and neurological

explanation and the connection between neurological and psychological terms is not specified. We find such statements as:

> ...reverberatory activity, the continuous firing of a neuron circuit or network over a period of time. This activity appears to account for ability to comprehend sentences [Smith, 1962, p.24].

> Substrata factors are thought of as neurological memory systems composed of smaller subsystems of the brain containing various kinds of information, such as audio, visual and kinesthetic associations which in a cultural milieu bestow a sense of reality upon symbolically represented thought units [Holmes & Singer, 1966, p.3].

These statements sound more sophisticated than "finding the main idea of a paragraph" but the use of technical language suggests more precise knowledge than is possessed. It must be shown exactly how a "reverberatory circuit" accounts for sentence comprehension, and how "neurological" sensory systems "...bestow a sense of reality upon symbolically represented thought units," and what is the set of symbols that represent thought units in the brain etc. Fodor (1968) has argued that adequate psychological theories are prerequisites to neurological explanations of psychological phenomena.

3. None of the models is based upon an adequate description of the objects of the comprehension process i.e., a theory of language.

4. Some of the models don't seem to generate any testable hypotheses. Sparks and Mitzel (1966) have criticized the Holmes and Singer (1966) model on this basis.

The models approach appears worth pursuing but better and more sophisticated models are needed. Models based upon a more careful look at language and the range of phenomena to be explained, and models with a better idea of the requirements a model must meet in order to provide insight into a phenomena, are needed before this approach can bear fruit.

Conclusion: Approaches to Comprehension

Taken all together the great deal of effort expended in the various approaches to comprehension have produced very little knowledge of reading comprehension that is commensurate with the effort expended. As Spache (1962) points out we still do not know "...1) exactly what thinking processes operate in comprehension, 2) how may the reader's facility in each of these processes be measured, and 3) how can ability in these processes be improved in instruction [p.63]?" The accumulated knowledge of the comprehension process does not go much beyond Thorndike's description quoted at the beginning of this chapter.

Comprehension Research and Pedagogy

The assumption underlying most educational research is that the effectiveness and therefore validity of pedagogical techniques is at least in part dependent upon the extent to which techniques capitalize upon the actual psychological processes that students utilize in learning. All other things being equal, the more effectively the pedagogical technique takes advantage of the student's actual learning process the more effectively the student will learn.[3]

In the case of reading comprehension, since very little is known about the comprehension process, the techniques for teaching comprehension are of unknown validity.

As a result of this lack of knowledge of the process, the techniques for teaching comprehension tend to be global and superficial without any real methods of focusing upon specific areas of student difficulty. In much of the material for teaching comprehension the student is presented a passage to read and then asked a series of questions on what he has read. These questions are stated in terms of the previously discussed skills, i.e. find the main idea, recall a detail, etc. If the student answers the question correctly, he then repeats the procedure on a different and more difficult passage. If he selects a wrong answer, he repeats the process on another passage of equal or lower difficulty. There is no attempt when the student makes an error to show him why he made the error or how to go about finding the correct answer.

In general when there is a lack of good description of psychological processes, the procedures for measuring performance are often used as a model for pedagogy. Consequently teaching procedures do not go very much beyond duplicating testing procedures. There is nothing in this procedure which allows for the correction of errors. For example the workbooks used in teaching vocabulary contain exercises that are almost identical to items on vocabulary tests. A further example is the use of the Cloze procedure to teach comprehension.

---

[3] If this assumption is incorrect then there is no reason to base teaching techniques on the type of educational and psychological research that aims to shed light on basic learning processes. And consequently this research would have no practical instructional implications.

In some cases there are attempts to go beyond testing procedures and to provide students with rules of thumb as principles. However these usually prove to be superficial and weak. For example in the case mentioned earlier of the skill of "distinguishing fact from opinion" the rule of thumb that is given to the students is a very superficial rule that doesn't hold particularly well (Schell, 1967).

Thus the pedagogical techniques for teaching comprehension leave a great deal to be desired. They boil down to practice in reading and testing and in some cases to superficial rules of thumb that don't seem to work very well. This crude superficiality of the techniques for teaching comprehension is all that can be expected when so little is known about the comprehension process. In fact it is difficult to justify all the voluminous research in reading comprehension if the only practical results, beyond readability formulas, are the present techniques for teaching comprehension. Indeed the meagre results of this line of research are the best argument for a new approach, one yielding increased knowledge of the comprehension process. This type of research would produce improved techniques for teaching comprehension.

In the next section some important reasons why the comprehension process has been resistant to explanation will be considered. These reasons will help in deciding upon the direction new work in comprehension should take in order to be more productive than past research.

Why the Comprehension Process has been Resistant to Explanation

In searching for the reason for the lack of success of past research efforts in comprehension, it would be misguided to place the blame upon the poor methodology that this research has exhibited. It could thus be assumed that methodologically more sophisticated research would produce meaningful results. While it is true that many of the past studies, as in any field of research, have had severe methodological difficulties (Cleland, 1964), improving methodology alone will not produce major breakthroughs in knowledge of the reading comprehension process, as the methodologically improved research over the past decade or so demonstrates.

The reason for the lack of progress in comprehension research goes much deeper than poor research methodology. The current lack of descriptions of the mental processes involved in reading comprehension render it difficult if not impossible to establish adequate behavioral criteria for successful comprehension. This in turn is due to the fact that comprehension is a complex and covert mental process which along with other cognitive processes takes place without any overt behavior being produced.[4] Thus the behavior that is measured in research, i.e. answers to questions on comprehension tests, may be only indirectly if at all related to the comprehension process. Furthermore, it is difficult to distinguish, in anything but an arbitrary way, between behavior that reflects the comprehension process from behavior that reflects other psychological processes -- such as motivation, memory, attitude, attention, personality, etc. To put the problem simply, it is almost impossible to conduct fruitful empirical research when there is a lack of knowledge of which behaviors provide relevant measures of the process under investigation.

Comprehension test developers have not really faced up to this problem. As a consequence there has been a great proliferation of comprehension tests measuring many different behaviors with no agreement among tests as to what the relevant behaviors are or the best ways to measure them.[5] Fodor (1968) points out this situation as typical of disciplines weak in theory. Spache (1962) describes the situation accurately:

---

[4] It is true that eye movements and pupil activity can be observed during comprehension. But these phenomena are difficult to interpret. Also, there are no doubt neurological changes that take place during comprehension, but these are not easily monitored. Even when they can be monitored they also have resisted interpretation. The reason, as Fodor (1968) points out is that neurological explanations of behavior are dependent upon adequate psychological explanations which of course are lacking in the case of comprehension.

[5] The difficulty in defining the appropriate behavior as a criterion for comprehension can be contrasted to the word recognition process in which there is a clear cut and easily definable behavioral criterion for successful performance i.e., correct oral pronunciation. For this reason word recognition appears to be more amenable to empirical research.

> Test makers are prone to define comprehension in terms of the types of questions included in their instruments. If six types of questions follow each reading selection, it is assumed that six aspects of comprehension or six types of thinking are being measured [p.61].

This problem is not unique to psychology, Kuhn (1962) points out, it is characteristic of the early stages of scientific development:

> In the absence of a paradigm (theory), all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant. As a result, early fact gathering is far more a nearly random activity than the one that subsequent scientific development makes familiar [p.15]. (Parentheses mine)

That comprehension research is at this "early fact gathering" stage with the concomitant inability to separate relevant from irrelevant facts is evidenced by the proliferation of comprehension skills and tests to measure them and the inability to distinguish comprehension from other psychological processes. The problem is not that more facts are needed but that it is impossible at present to give a coherent explanation of the facts that already exist.[6]

The Need to Base Comprehension Research on Theory

Clearly an important step that must be taken to help remedy the state of affairs that exists in comprehension research is to base future research on theory. Theory based research has the advantage of helping to provide a principled way of separating relevant from irrelevant facts, of determining appropriate behavioral criteria, and suggesting important hypotheses that can be subjected to empirical research. Theory allows empirical research to be conducted on a rational and systematic basis.

---

[6] The point of scientific inquiry is to provide through theories an explanation of phenomena. Facts or data are important only as they bear upon the truth of falsity of theory. They only make sense in light of theory. In the absence of theory facts are of minor importance to scientific inquiry.

Unfortunately the term "theory" has been used rather loosely by reading researchers and some models and taxonomies have been mistakenly called theories. An examination of the criteria for scientific theories (Schutz, 1962) will quickly convince the reader that the term theory does not apply to most models and taxonomies.

Two Types of Theories: Competence and Performance

There are two types of theories that are relevant to the study of psychological processes such as reading comprehension. These are theories of competence and theories of performance, (Fodor and Garrett, 1966). A competence theory is a theory of what someone knows while a performance theory is a theory of how someone uses this knowledge. In the case of reading comprehension a competence theory describes what is comprehended, i.e. the language and thoughts and ideas, while a performance theory describes how the comprehension process works. In other words a competence theory is a description of the object of the process while a performance theory describes the actual process, i.e. mental manipulations.

An example will help clarify the competence/performance distinction. In the reading skill of learning to recognize words, the object of the process, i.e. what it is that is learned, is the relationship between printed words and the pronunciation of these words. A competence theory would in effect be a description of the relationship between the English spelling and sound system. The actual psychological processes that readers employ in decoding English spellings into correct pronunciations is described by performance theory.

The Need to Limit the Scope of Comprehension Research

It appears to be necessary to limit the scope of comprehension research to those aspects of comprehension for which there exist adequate competence theories, i.e. descriptions of the tool employed in the process, namely language. This restriction is necessary because the development of competence theories is prerequisite to the development of performance theories. It is necessary to have a description of what it is that is learned before it is possible to conduct fruitful research into the learning processes. An example will clarify this point. In the putative comprehension skill of "fusing new and old ideas" it seems pretty clear that before empirical study of this skill can be successfully undertaken, it will be necessary to define ideas, to develop a perfectly explicit method of counting ideas, to develop an unequivocal procedure for analyzing text into its component ideas etc. An adequate competence theory of ideas would, if it existed, provide this information. This theory would then provide a principled way of counting ideas as well as help to distinguish between relevant and irrelevant behavior in the development of behavioral criteria for measuring the fusion of ideas in comprehension. A competence theory of ideas would also suggest testable hypotheses about the fusion process which would eventually result through empirical research in a performance theory of the process by which readers "fuse new and old ideas." Without a competence theory of ideas it is almost impossible to know where to begin.

The lack of a theory of ideas has resulted in research in which either ideas are so narrowly defined as to lose their resemblance to the original concept of an idea or in research in which such global and vague definitions of ideas are employed that the results are virtually uninterpretable. In both cases not much light has been shed on the nature of ideas or upon the ways in which readers learn to "fuse new and old ideas" in comprehending written matter.

Many of the attempts to understand the comprehension process have been less than successful precisely because of the lack of this competence theory of the object of the process upon which to base research. Without this competence theory the research has been reduced to almost "random fact gathering."

Furthermore, it should be noted that basing comprehension research on available linguistic competence theory means that this research at present must be limited to attempts at understanding the processes involved in comprehending the literal meaning of sentences. This limitation is necessary because linguistic competence theory has only advanced to the stage of dealing with the literal meaning of sentences, and there is little, if any, theory that deals with anything beyond the sentence level. Moreover, the literal interpretation of sentences is a good starting point for comprehension research because most of the more global comprehension skills discussed earlier assume a literal interpretation as a prerequisite. As Huey (1908) points out, "Language begins with the sentence and this is the unit of language everywhere [p.123]."

Deep Structure as a Description of the Object of the Comprehension Process

Given that reading comprehension research should be limited to those aspects of comprehension for which a theory of language currently exists, the question to ask is: for what aspects of comprehension do there exist competence theories? It seems clear that one aspect of what is understood in comprehension is the underlying structural relationships of sentences. For example in the sentences:

(a) John is eager to please

(b) John is easy to please

The reader must understand that in (a) John is the one who is doing the pleasing and the person that is pleased is unspecified, while in (b) the reader must understand that it is John who is pleased and the person doing the pleasing is unspecified. This information is necessary to understand a sentence. To comprehend a sentence the reader must understand the underlying structural relationships i.e., the logical subject and logical object of the sentence. Thus one aspect of the object of the comprehension process is the underlying structural relations of sentences. Fortunately there does exist a competence theory that describes these relationships. This theory is generative or transformational grammar (henceforth T.G.), as developed by Noam Chomsky (1957, 1965). In T.G. it is the deep structure that captures these underlying structural relationships.[7] (See Jacobs & Rosenbaum, 1968; and Langacker, 1967 for readable descriptions of transformational grammar). T.G. thus provides the competence theory of the object of the comprehension process upon which the research reported here is based.

But before describing this study it is necessary to ask whether or not the deep structure of sentences as described by T.G. is psychologically real? In other words is the deep structure actually employed in language comprehension? The studies bearing on this question will be discussed in the next section.

---

[7] These relationships are deep or underlying in the sense that they are not usually given in the surface structure of sentences. This can be seen in the example above in which (a) and (b) both have the same surface structure but different deep structures.

The Psychological Reality of Deep Structure

A number of studies have attempted and have been without exception successful in demonstrating the psychological reality of deep structure. They have done this by showing that deep structure differences with surface structure held constant are consistently reflected in recall, recognition, comprehension and learning of sentences. (Bever, Mehler & Carey, 1967; Blumenthal, 1967; Blumenthal & Boakes, 1967; Clark, 1969; Davidson & Dollinger, 1969; Levin & Wanat, 1967; Levelt, in press; Mehler & Carey, 1968; Wanner, 1968). A representative sample of these studies is described below.

Blumenthal (1967) compared subjects' recall of two types of sentences. Recall was aided by a prompt word taken from the sentence. Both types of sentences had the same surface structure but they differed in their deep structures. The first type was a standard passive, e.g., "Gloves were made by tailors." In the second type of sentence the by-phrase was replaced with a non-agent adverbial by-phrase, e.g. "Gloves were made by hand." In the first type "tailors" is the deep structure logical subject of the whole sentence, e.g., tailors made gloves. In the second sentence "hand" is a verb modifier and does not relate to the whole sentence in the deep structure.

The results of the experiment support the psychological reality of deep structure. Blumenthal found that the first type of sentence was recalled more easily than the second type. He also found an interaction between prompt words and sentence types. When the initial noun was the prompt word, there were no differences in the ability of subjects to recall the two types of sentences. But when the final noun was the prompt word, the standard passive, e.g., "Gloves were made by tailors," was more easily recalled than the second type, e.g., "Gloves were made by hand." These results could be predicted by the differences in the deep structure relations in the sentences. In the first sentence, "tailors" is the deep structure subject of the whole sentence, while in the second sentence, "hand" is related only to a part of the sentence as an adverbial.

In a subsequent study in which different types of sentences were used and a more difficult recall task employed, Blumenthal and Boakes (1967) provided further confirmation of the psychological reality of deep structure. As in the previous experiment, they used sentences with the same surface structure but with different deep structures, e.g., "John is easy to please," and "John is eager to please." In the first sentence "John" is the object of the sentence and the adjective "easy" is related to the whole sentence as a modifier, i.e. a paraphrase of the sentence could read "For someone to please John is easy." In the second sentence "John" is the subject of the sentence and the adjective "eager" is related to only part of the sentence as a noun modifier. The results were the same as in the earlier study. The sentence modifier, e.g., "easy", when given as a prompt word, produced better recall of its sentence than the noun modifier, e.g., "eager", did for its sentence.

Further evidence is provided by Levin and Wanat (1967). They measured Eye-Voice-Span (the amount that a Ss' eyes are ahead of his voice when he is reading aloud) of subjects reading passive sentences that had the same surface structure but different deep structures. Two types of passive sentences were used. One type was the standard passive, e.g., "His brother was beaten up by the gang." In the other type the underlying subject of the sentence was deleted, e.g., "His brother was beaten up by the park." They found a larger EVS for the first sentence type than for the second sentence type. Thus deep structure differences produce differences in Eye-Voice-Span.

Finally, a study by Bever, Mehler and Carey (1967) demonstrated that eye fixation patterns during reading were also influenced by differences in surface and deep structures of sentences. They used ambiguous sentences, i.e., sentences that can be interpreted in more than one way. Two types of ambiguities were used in this study -- surface structure and deep structure ambiguity. Surface structure ambiguities are those in which ambiguity is contained in the grouping of the words of the sentence. An example of surface structure ambiguity is "They gave her dog candies." This sentence has two interpretations: They (gave (her dog) candies), i.e., candies were given to her dog, and They (gave her (dog candies)), i.e. dog candies were given to her. A deep structure ambiguity is in the underlying logical relations of the sentence. "The shooting of the hunters..." is ambiguous at the deep structure level. It has two interpretations: (The shooting (of (the hunters))) i.e., the hunters were shooting, and (The shooting (of (the hunters))) i.e., somebody was shooting the hunters. Notice that while the deep structures differ, the surface structures as denoted by the parentheses remain the same.

The eye fixations of the subjects were recorded as they read stories five sentences long in which the ambiguous sentence was the fourth one. The story context made one interpretation of the ambiguous sentence highly probable. The eye fixation patterns were compared for the two interpretations of the ambiguous sentence. They found that deep structure as well as surface structure differences were reflected in fixation patterns.

These studies taken all together provide ample evidence for the psychological reality of deep structure. Thus T.G. provides a psychologically real description of the object of the comprehension process, i.e., the deep structure. This theory provides the basis for the study of reading comprehension reported in this dissertation.

CHAPTER II

An Investigation of Deep Structure and Reading Comprehension

The Purpose of the Study

The study described in this paper is concerned with the relevance of a specific linguistic concept -- deep structure -- to reading comprehension. Ths study has two concerns: first to investigate the relationship between children's skill at recovering the deep structure of sentences and their reading comprehension skill, and second to investigate one strategy that students might use in recovering the deep structure of sentences and to determine the relationship of this strategy to reading comprehension.

Hypotheses and Assumptions of the Study

Hypothesis 1: Children's skill at recovering the deep structure of sentences is positively related to reading comprehension skill.

This hypothesis rests upon the following assumptions:

1. Comprehending a sentence minimally involves recovering the deep structure of that sentence.
2. Understanding sentences is positively related to understanding connected discourse.

An objection could be raised here. It might be argued that since recovery of deep structure is almost by definition an aspect of sentence comprehension, then reading comprehension of connected discourse must involve recovery of deep structure because reading comprehension involves understanding sentences. According to this line of reasoning hypothesis 1 is a tautology and bound to be confirmed by the data. It is, on the contrary, not at all clear that this hypothesis will produce positive results, for the following two reasons:

1. Recovery of deep structure might be such a fundamental skill that children all employ it equally well. Thus there would be no individual differences in this skill.
2. There may be differences in children's skill at recovering deep structure, but this skill is totally submerged by other factors such as word knowledge, past experience, motivation, attention etc. in accounting for reading comprehension.

For either of these two reasons correct recovery of deep structure could play little if any role in reading comprehension and consequently hypothesis 1 would not be confirmed.

Hypothesis 2: Children's skill in making a lexical analysis of the main verb of sentences is positively related to their reading comprehension skill.

This is a strategy that children may use in recovering the deep structure of sentences i.e. a lexical analysis of main verbs, has been suggested by Fodor, Garrett, and Bever (1968). They suggest that

"The exploitation of the lexical analysis of the main verb of a sentence is a central heuristic in the strategy Ss use to recover its deep structure [p.459]." This strategy involves analysis of the main verb of the sentence to determine the possible deep structures with which it is compatible. This information is formalized in the grammar by the strict subcategorization rules which classify verbs according to the deep structure configurations with which they are compatible (Chomsky, 1965). For example, some classes of verbs can take more deep structures than others. Thus, the verb "hit" is a pure transitive and can only take one deep structure as in (a),

(a) The boy hit the girl.

The verb "know", however, can take two deep structures, a transitive, as in the sentence (b),

(b) The boy knows the girl,

and a complement as in the sentence (c),

(c) The boy knows that John is smart.

If the strategy for recovery of the deep structure involves the information contained in the strict subcategorization rules, then all other things being equal, verbs such as "know" should require more complex processing than verbs such as "hit" because the former entertains more possible deep structures than the latter.

Fodor, Garrett, and Bever (1968) have provided evidence in support of this strategy by demonstrating that sentences containing verbs that can take either a complement or transitive deep structure, e.g. The box the man the child knew carried was empty, are more difficult and take longer to paraphrase than the same sentence with knew replaced by a main verb that can take only a transitive deep structure e.g. The box the man the child met carried was empty.

Hypothesis 2 rests on assumption 1 and 2 previously stated plus the following assumption:

3. A lexical analysis of the main verb is a part of the strategy that is used in the recovery of deep structure.

In the next chapter the measurement of the variables of these two hypotheses will be described.

CHAPTER III

The Measurement of Variables Described

Measurement of the Variables for Hypothesis 1

Hypothesis 1: Children's skill at recovering the deep structure of sentences is positively related to reading comprehension skill.

Independent Variable: Skill at recovering the deep structure of sentences.

This variable is measured by a 25 item test developed by the author for this study. The items on this test require Ss to choose the one sentence out of three that is not a paraphrase of the other two. Some examples follow:[1]

1. a. The boy gave the book to the girl.
   b. The book was given the girl by the boy.
   *c. The book was given to the boy by the girl.
2. *a. What the boy would like is for the girl to leave.
   b. For the boy to leave is what the girl would like.
   c. What the girl would like is for the boy to leave.
3. a. He painted the red house.
   *b. He painted the house red.
   c. He painted the house that was red.

---

[1] The asterisk indicates the sentence with a deep structure different from the other two sentences.

4. a. The girl asked the boy when to leave.
   b. The girl asked the boy when she should leave.
   *c. The girl asked the boy when he should leave.

5. *a. The girl who the boy hit fell down.
   b. The boy the girl hit fell down.
   c. The boy who the girl hit fell down.

The complete test is reproduced in Appendix A. This test will henceforth be called the Deep Structure Recovery Test (D.S.R.T.).

A close look at the test reveals the characteristics of the items. In each item there are two sentences that have the same deep structure and the same meaning. These two sentences have different surface structures due to the application of one or more extra transformations to one of them. The third sentence has a surface structure that is the same or similar to one of the other two sentences, but it has a different deep structure and consequently a different meaning from the other two sentences. An example will clarify this description.

In the sample item,

a. The boy hit the girl.
b. The girl was hit by the boy.
*c. The boy was hit by the girl.

a and b both have the same deep structure but have different surface structures in that b has had the passive transformation applied to it. C has the same surface structure as b but has a different deep structure than b because in c "girl" is the logical subject while in b and a "girl" is the logical object. The same reversal of deep structure relationships hold for "boy." In addition the same content words, i.e. nouns and verbs, are used in all three sentences of each item. Overall in each item the sentence with the different deep structure from the other two sentences is the same in almost all characteristics as at least one of the other sentences except that it has a different deep structure.

Validity of the D.S.R.T. The validity of the D.S.R.T., as a measure of the ability to recover deep structure, rests on the following assumptions:

1. Sentences that are paraphrases of each other have the same deep structure.
2. It is necessary to recover the deep structure in order to determine whether or not sentences are paraphrases of one another.
3. Choosing the correct answer on items of the D.S.R.T. indicates that the deep structures of at least two of the three sentences in each item have been recovered.
4. The D.S.R.T. is measuring the recovery of deep structure. It is not measuring certain other variables such as word recognition and word knowledge that are associated with reading comprehension.

5. The statistical properties of the test are adequate.

Each of these assumptions will be discussed below.

Assumption 1: Sentences that are paraphrases of each other have the same deep structure.

This is simply a fact of language as it is described by T.G. If T.G. is wrong about this fact, then the assumption is false. If T.G. is correct, then the assumption is correct. Until T.G. is proven to be wrong as a theory of language competence or the theory changes so that this assumption no longer is a fact of language, it seems reasonable to accept the correctness of assumption 1.

Assumption 2: It is necessary to recover the deep structure in order to determine whether or not sentences are paraphrases of one another.

The evidence presented in Chapter I that recovery of deep structure is necessary for understanding a sentence plus assumption 1 offer strong support for this assumption. The line of reasoning supporting this assumption is that understanding a sentence involves recovery of the deep structure of that sentence. If two sentences have the same deep structure i.e., have the same meaning, an S must recover the deep structure of both sentences in order to determine if they are paraphrases of each other.

Assumption 3: Choosing the correct answers on items of the D.S.R.T. indicates that the deep structures of at least two of the three sentences in each item have been recovered.

In order to show the validity of this assumption it is necessary to rule out other alternative ways of getting the correct answer or ways that do not require recovering deep structure of the sentences. The first alternative is that Ss get the correct answers by random guessing. This alternative can be ruled out by looking at the mean total scores on the test for the sample studied and comparing this score to the score that is possible by random guessing. Since there are three choices per item, Ss have 1/3 chance of getting the right answer on each item by guessing. There are 25 items on the test. This makes 8.33 a chance score on the total test assuming that all choices are equally probable. The mean score of the sample studied was found to be 18.575. This is sufficiently different from 8.33 to rule out random guessing as an alternative.

A second related alternative is that Ss choose some set position i.e. first, second, or third sentence in doing the test. This alternative was controlled for by attempting to balance the position of the correct answer over all 25 items. An error in this procedure produced the following partially unbalanced distribution of correct answers over positions:

(a) First position -- 8 items

(b) Second position -- 7 items

(c) Third position -- 10 items

However, even if Ss adopted an optimal position strategy and chose the third position their scores would be 10 correct. This score is sufficiently different from the Ss' mean score of 18.575 to rule out the alternative that Ss choose the answer by its position.

A third alternative is that the length of the sentence influences the Ss' choice. There are a few different possibilities here. Ss could adopt a strategy of choosing the shortest sentence in each item or the longest sentence without reading the sentences at all. Or they could actually read the sentences and choose the longest sentence because it is the most complex. Thus the complexity of the sentence somehow determines performance on these items rather than skill in recovering deep structure.

A look at the items and their sentence length appears to rule out this alternative. In one item (no.2) all the sentences are the same length. In two items (nos. 19, 24) the shortest sentence is the correct answer. In two items (nos. 7, 14) the intermediate length sentence is the correct answer. In three items (nos. 1, 13, 17) the longest sentence is the correct answer. In five items (nos. 3, 6, 8, 16, 20) the correct answer is one of two sentences both of which are the same length but shorter than the third sentence. And finally in twelve items (nos. 4, 5, 9, 10, 11, 12, 15, 18, 21, 22, 23, 25) the correct answer is one of two sentences of the same length both of which are longer than the third sentence. There is no plausible strategy that Ss could adopt that would result in a score close to the 18.575 average found. For example an optimal strategy of choosing the longest sentence and in the cases where there are two longest sentences chose one randomly would only result in 9 or 10 correct. It appears that the sentence length alternative can be safely ruled out.

A fourth alternative strategy that Ss might adopt for choosing the correct answer without recovering the deep structure of sentences in the items is to use some superficial characteristic of sentences. For example in item #13:

13. a. The boy begged the girl to tell the truth.
b. What the boy begged the girl to do was tell the truth.
*c. To tell the truth was what the girl begged the boy to do.

Ss might adopt a strategy of comparing the phrases underlined here and determine that c is the correct answer because the phrase is reversed. Or Ss might look at the first occurrence of the nouns "boy" and "girl" and choose a sentence in which the first occurrence of "boy" or "girl" is different from the other two sentences. Or Ss could use the same strategy with the second occurrence of "boy" or "girl." All these strategies and other similar strategies do not involve recovery of the deep structure of the sentences, in fact they hardly even require Ss to read the sentence at all.

There are 8 items out of the 25 in which the use of superficial clues will yield the correct answer. These items, with their superficial clues underlined, are presented in Table 1 along with their item difficulty.

In the remaining 17 items of the D.S.R.T. (nos. 1, 3, 4, 6, 7, 8, 9, 10, 11, 12, 14, 16, 19, 20, 21, 23 & 25), there are no superficial clues that Ss could use in choosing the correct answer rather than recovering the deep structure of the sentences. Thus the superficial clue strategy alternative can be ruled out for these 17 items. These items will be referred to as the No Sup. Clues items.

In order to examine Ss' use of the superficial clue strategy in the 8 items in Table 1 this strategy is actually possible, it is necessary to compare Ss' performance on the 8 items which contain superficial clues (Sup. Clues) with Ss' performance on the 17 items that do not contain superficial clues (No Sup. Clues). A look at the item difficulties in Table 1 reveals that 6 out of the 8 items are quite easy. Most Ss chose the correct answer on these items. A t test between the average scores on the Sup. Clues Items and the average scores on the No Sup. Clues items reveals that the Sup. Clues items were significantly ($p < .001$) easier than the No Sup. Clues items. This finding would tend to support the conclusion that Ss were using superficial clues on the Sup. Clues items. Ss got more

TABLE 1

ITEMS WITH SUPERFICIAL CLUES

(Superficial Clues Underlined)

Item #		N=103	Item Difficulty
2.	*a.	What the boy would like is for the girl to leave.	
	b.	For the boy to leave is what the girl would like.	.709
	c.	What the girl would like is for the boy to leave.	
5.	*a.	The girl who the boy hit fell down.	
	b.	The boy the girl hit fell down.	.495
	c.	The boy who the girl hit fell down.	
13.	a.	The boy begged the girl to tell the truth.	
	b.	What the boy begged the girl to do was tell the truth.	.922
	*c.	To tell the truth was what the girl begged the boy to do.	
15.	a.	The girl taught the boy to use a pencil.	
	*b.	What the boy taught the girl was to use a pencil.	.903
	c.	To use a pencil was what the girl taught the boy.	
17.	a.	The girl ordered the boy to sit down.	
	b.	What the girl ordered the boy to do was sit down.	.961
	*c.	To sit down was what the boy ordered the girl to do.	
18.	*a.	The girl who the boy knew went home.	
	b.	The boy the girl knew went home.	.631
	c.	The boy who the girl knew went home.	
22.	a.	The boy sees that the girl is tall.	
	b.	That the girl is tall is seen by the boy.	.825
	*c.	What the girl sees is that the boy is tall.	
24.	a.	What the girl wants is for the boy to find the ball.	
	*b.	The boy wants the girl to find the ball.	.874
	c.	For the boy to find the ball is what the girl wants.	

right answers on these items than on the No Sup. Clues items where they had to depend upon the recovery of deep structure exclusively.

Further evidence bearing on this conclusion can be found in the correlations between the Sup. Clues items and the No Sup. Clues items and the dependent variable of this study, reading comprehension as measured by a Cloze test and the Reading Subtest of the Metropolitan Achievement Test. (See discussion of the dependent variable on pp. 57-61 for the rationale for these measures). This data is presented in Table 2.

TABLE 2

Correlations between Sup. Clues and No Sup. Clues Items and Reading Comprehension (N=87)

Variable	Reading Comprehension	
	Cloze	MAT Reading
Sup. Clues	.536**	.393**
No Sup. Clues	.719**	.441**

$^{**}p < .01$

This data shows that the Sup. Clues items correlate significantly with both measures of the dependent variable. These correlations are higher for the No Sup. Clues items than for the Sup. Clues items and this difference is significant at the .05 level for the Cloze criterion but nonsignificant for the M.A.T. Reading test.

The similar correlations between the Sup. Clues and the No Sup. Clues items and the criterion measures suggest that the superficial clue strategy is used very little. The lowered correlation for the Sup. Clues items can be explained by a ceiling effect in which nearly 30% of the Ss got a perfect score of 8 on the test. This ceiling effect would of course tend to lower the correlation. The evidence presented thus far is not sufficient to rule out the use of superficial clues for the Sup. Clues items. However the effect on the results of the study is not important because a complete analysis with the No Sup. Clues items omitted produced essentially the same results as the analysis which included these items. Thus the inclusion of the 8 Sup. Clues items has no effect on the results or conclusions of this study. In this study, then, only the results from the total D.S.R.T. will be reported here.

Assumption 4: The D.S.R.T. is not measuring certain other variables associated with reading comprehension in addition to skill at recovering deep structure.

In order to make the D.S.R.T. a clean measure of the ability to recover deep structure, it is necessary to make sure that the D.S.R.T. is not measuring other variables that are associated with reading comprehension. If this is not done and the D.S.R.T. turns out to correlate with reading comprehension, then it will be hard to tell whether this correlation is due to skill at recovering deep structure or to some other variable that is being measured by the D.S.R.T. which correlates with comprehension.

One variable which the D.S.R.T. could be measuring and which is also related to reading comprehension is word recognition skill. It is possible that the Ss are making errors on the D.S.R.T. not because they can't recover deep structure but because they cannot recognize the words in the D.S.R.T. This possibility was ruled out in two ways: (a) The test was composed of a small set of 78 different words from a vocabulary for the primary grades developed for remedial work by Durrell (1956). Since the sample for the study was fifth grade students, it seems reasonable to assume that the words on the D.S.R.T. were easy to recognize. (b) In order to be absolutely certain that the Ss for this study are able to recognize the words on the D.S.R.T. the following procedure was employed. Ss were rank ordered on the basis of word recognition ability as measured by the Word Discrimination Subtest of the Metropolitan Achievement Test. Then beginning with the Ss with the lowest scores on Word Discrimination Test, Ss were tested individually on their ability to recognize the 78 individual words on the D.S.R.T. The Ss were required to pronounce all 78 words. Every S with a Word Discrimination score below 5.0 was tested as it was assumed that the rest of the Ss who had scores higher than 5.0 on the Word Discrimination test would be able to recognize the words. Only 9 Ss made errors on the 78 words. Seven made only 1 error, 1 made

2 errors and 1 made 4 errors. Ss were eliminated if they made mistakes on any item of the D.S.R.T. which contained words which they had failed to recognize on the oral pronunciation test. Using this criterion, 7 of the 9 Ss were eliminated from the sample. Therefore the alternative that the D.S.R.T. measures the ability to recognize words can be ruled out because words were easy to recognize and those Ss with trouble were eliminated.

Another variable which the D.S.R.T. could be measuring is Ss' knowledge of word meanings. Ss' errors on the D.S.R.T. may be due to their lack of knowledge of the meanings of the words on the test rather than their ability to recover deep structure. Since word knowledge is generally correlated with reading comprehension, a positive correlation between the D.S.R.T. and reading comprehension may be due to Ss' knowledge of the word meanings on the test rather than their ability to recover deep structure.

There is reason to believe that the knowledge of the meanings of the words on the D.S.R.T. is not a major factor in performance on this test for the sample studied. The words used are at primary grade level (Durrell, 1956) and are on the list of easy words used in the Dale-Chall readability formula. In addition, from a simple examination of the words on the test, it appears that they are quite easy and common. Finally, the sample studied were fifth grade students who were on the average above grade level on the Word Knowledge subtest of the M.A.T.

and who had an average I.Q. of 117. Given this information about the difficulty of the words on the test and the ability of the Ss, it seems highly unlikely that these Ss do not know the meanings of the words on the D.S.R.T.

Assumption 4 is tenable in that the D.S.R.T. does not appear to be measuring word recognition skill or knowledge of word meanings.

Assumption 5: The statistical properties of the test are adequate.

On the basis of data collected on 103 fifth grade students the following statistical properties of the test were found:[2]

1. The item difficulties ranged from .282 to .96.
2. The internal validity as measured by a biserial correlation between the test items and the total score showed items with internal validities ranging from .358 to .842.
3. The Kuder-Richardson 20 reliability of the total D.S.R.T. is .7964.

See Appendix B for the Complete item statistics of the test. Table 3 presents the basic statistics for the total D.S.R.T.

---

[2] This same population was used to test the hypotheses of the study. Thus there was no independent statistical validation of the D.S.R.T.

TABLE 3

Basic Statistics of the Total D.S.R.T.
(Sum of items correct) N=103

Variable	$\bar{X}$	Median	Range	S.D.	Skewness
D.S.R.T. (total 25 items)	18.575	20.00	3.25	4.258	-0.767

The following conclusions can be drawn from this data concerning the statistical properties of the test:

1. The reliability of the test is acceptable but slightly low.

2. The test is relatively easy with an average total score of 18,575 out of 25 items. A number of items were quite easy, with item difficulties over .7.

3. The items for the most part are good discriminators with most of the items exhibiting biserial correlations well over .5.

Taken together this data makes assumption 5 tenable. The statistical properties of the test are adequate, (Guilford, 1951).

In summary, all the assumptions upon which the validity of the D.S.R.T. rests appear to be sound. The D.S.R.T. by meeting these five assumptions is a valid test of skill in recovering the deep structure of sentences. Consequently it has construct validity; A relationship between the measurement and the underlying skill, i.e. ability to recover deep structure, has been established.

The Dependent Variable: Reading Comprehension Skill

As discussed in Chapter 1 the measurement of reading comprehension skill presents a difficult problem because not much is known about the reading comprehension process. As a result, in measuring comprehension skills it is difficult to identify the relevant behaviors that must be measured as well as to distinguish the measurement of comprehension skills from other student skills. Until a theory of the comprehension process is developed there will be no completely satisfactory solution to this problem. It is possible, however, to distinguish better measures of comprehension from poorer measures of it even though they all lack construct validity.

The best measure currently available for the measurement of reading comprehension is the Cloze test as discussed in Chapter 1. The reasons that the Cloze provides a better measure of comprehension than the traditional standardized test are discussed below:

1. As discussed in Chapter 1 the Cloze test measures fewer extraneous aspects of student functioning than traditional tests do.
2. Cloze tests provide a measure of comprehension as this process is in progress, not after the process is completed as is the

case with traditional tests. In this way the Cloze test is a more direct measure of the comprehension process than traditional tests.

3. Reliabilities of Cloze tests (Weaver & Kingston, 1963) compare favorably with reliabilities of traditional tests and meet the standards for good test construction. For example Weaver & Kingston (1963) found reliabilities in the high .80's and low .90's.

The concurrent validity of Cloze tests as a measure of comprehension has been demonstrated by a number of studies which have shown correlations between the Cloze and traditional measures of comprehension ranging from .50 to .80. (See Potter, 1968 for a summary of these studies). In addition Cloze tests have been used in readability studies and have been shown to be as good or better measures of comprehension than traditional tests (Bormuth, 1966; Potter, 1968).

The evidence on the concurrent validity of Cloze tests indicates that they are equally good measures of reading comprehension as traditional tests, while the three points presented above indicate that Cloze tests are more direct measures of the comprehension process. Since this study is concerned with the comprehension process, the Cloze test is the major measure of reading comprehension in it.

The Cloze Tests of Comprehension Used in this Study. Three 150 word passages were selected from a set of passages calibrated on 475 college students by Miller and Coleman (1967). The passages come from The McCall Crabbs Standard Test Lessons in Reading (1961). The paragraphs are reproduced in Appendix C. The difficulties of the passages as determined in the Miller and Coleman study are presented in Table 4.

TABLE 4

Cloze Passage Calibration

Scoring Criterion	Cloze Test Passage 1.	Cloze Test Passage 2.	Cloze Test Passage 3.
A. % Correct Responses (Every 5th word deletion)	78.9	68.5	55.7
B. % Correct Responses (One word per passage deletion)	88.2	75.3	66.7
C. % Correct Responses (S guesses every word after seeing preceding word)	42.1	34.4	35.0
D. % New Correct Responses (Ss repeated C above)	38.8	42.1	43.0
E. Total number of Correct Responses on A.B.C. (2400 possible correct responses)	1389	1215	1107
F. McCall Crabbs grade level score for 90% comprehension	7.0	9.0	6.6
G. McCall Crabbs grade level score for 70% comprehension	6.7	6.0	6.0

All three passages were prepared with every fifth word deleted i.e., 30 deletions per passage. Passages 1 and 2 were administered in the standard way, Ss were simply instructed to fill in the blanks. For Passage 3 a slightly different procedure was used. Ss first read the intact passage. Then without the intact passage to refer to they were given the same passage as a Cloze test with every 5th word deleted. There were no time limits on the test. For each test the number of exact correct replacements was totaled. Table 5 presents the results of the test on 87 5th grade students plus the Rulon split half reliability (Guilford, 1951). All three tests have high reliabilities.

TABLE 5

Basic Statistics on Three Cloze Tests (N=87)

Variable	No. Items	$\bar{X}$	Median	S.D.	Range	Skewness	Rulon Reliability
Cloze Test #1	30	17.724	18.00	3.646	5-25	-0.701	.884
Cloze Test #2	30	16.138	16.00	4.243	5-27	-0.024	.924
Cloze Test #3	30	16.023	16.00	3.782	4-24	-0.347	.893

The three Cloze tests were converted to T scores and summed to form a composite Cloze measure of reading comprehension. This measure was used as the dependent variable of the study. Table 6 presents the intercorrelations of the three Cloze tests plus the composite test.

TABLE 6

Intercorrelations: Cloze Tests* (N=87)

Variable	1	2	3	4	$\bar{X}$	S.D.
1. Cloze #1	1.000	.497	.491	.805	17.724	3.646
2. Cloze #2		1.000	.565	.834	16.138	4.243
3. Cloze #3			1.000	.832	16.023	3.782
4. Composite Cloze (T score)					150.000	24.712

*All correlations are significant at the .01 level or less.

In addition to the Cloze test described above a traditional measure of reading comprehension was used. This test is the Reading subtest of the Metropolitan Achievement Test (M.A.T.). This test has all the inadequacies of traditional comprehension tests as described in Chapter 1. This test is only included to provide a basis of comparison with a traditional comprehension measure. The data on the M.A.T. Reading is presented in Table 7.

TABLE 7

Basic Statistics on M.A.T. Reading[a]

Variable	$\bar{X}$	Median	S.D.	Range	Skewness
M.A.T. Reading	4.984 (5.4)	5.2 (5.7)	1.112	2.2-6.5 (2.8-8.4)	-0.605

[a] Local norms (National norms in parentheses)

Control Variables

In testing a hypothesis in which a relationship between an independent variable and a dependent variable is hypothesized, it is necessary to rule out alternative explanations of the findings by controlling other variables that may be related to both the independent and dependent variable and thus account for any relationship found between the independent and dependent variable. For the hypothesis 1 of this study a finding of a positive relationship between the D.S.R.T. and the Cloze test may be due to some other third variable rather than the hypothesized variables. These third variables must be controlled and the relationship of D.S.R.T. with the Cloze test must be looked at with the effect of these variables removed. The three control variables for this study are:

1. Knowledge of word meanings as measured by the Word Knowledge subtest of the M.A.T. Elementary Battery Form B. In this test the word to be defined is presented in a brief sentence with one word underlined. The student is required to select from four choices the word which best completes the sentence. The word is usually a synonym for the underlined word.
2. Word recognition skill as measured by the Word Discrimination subtest of the M.A.T. Elementary Battery Form B. This test consists of sentences with a word missing. The student selects the proper word to complete the sentence from among several words differing in configuration and meaning.

3. Intelligence (I.Q.) as measured by the Lorge-Thorndike Intelligence Tests, Form A, Level 3, Verbal Battery. Part 1 requires a student to read a sentence with a word deleted and choose from among five alternatives the word that best completes the sentence. Part 2 requires the student to choose a word that is of the same category as a set of stimulus words. Part 3 requires the student to solve arithmetic word problems and select the correct answer. Part 4 requires the student to choose a synonym for a stimulus word from among five alternatives.

These three tests are typical of traditional ways of measuring word recognition, vocabulary and intelligence. In order to be useful variables, they should provide clean measures of what they purport to measure and nothing else. Unfortunately they appear to suffer as do comprehension tests from a lack of construct validity. Because of this lack there is an inability to choose the relevant behaviors to measure and these tests end up measuring more than their putative variables. They seem to be measuring some common skills. In the first place they all require the student to recognize the words on the test as a prerequisite to successful performance on the items. Thus to some unknown degree they all are measuring students' ability to recognize words. In the second place the Word Knowledge, Word Discrimination and the I.Q. tests are all measuring word meanings. Thirdly all three tests contain items that are quite similar to a Cloze test in the sense that deleted words from sentences must be identified. Thus they are measures of comprehension as well. Thus on the face of it they have much in common with each other as well as with the dependent variable of this study, the Cloze test. This interdependence can be seen in the intercorrelations of these variables as presented in Table 8.

TABLE 8

Intercorrelations: Dependent and Control Variables* (N=87)

Variable	1	2	3	4	5	$\bar{X}$	S.D.
1. Word Discrimination	1.000	.761	.585	.663	.542	4.714	0.886
2. Word Knowledge		1.000	.674	.740	.621	4.560	1.005
3. I.Q.			1.000	.587	.484	117.011	12.974
4. M.A.T. Reading				1.000	.528	4.984	1.112
5. Cloze Test					1.000	150.000	24.712

*All correlations significant at the .01 level or less.

The global nature of the control variables plus their similarity to the Cloze comprehension test should tend to make it more difficult to show a relationship between the D.S.R.T. and the Cloze test when these variables are controlled. Because they measure so much, these variables might be expected to account for much of the relationship between D.S.R.T. and Cloze comprehension if indeed a positive relationship is found.

The same type of argument can be raised in terms of the control variables as they relate to the traditional measure of comprehension i.e., the Reading subtest of the M.A.T. First the M.A.T. Reading subtest is measuring ability to recognize words since an S must be able to recognize the words before he can do the test. Second, 6 of the 44 items on the M.A.T. Reading test ask for the meanings of specific words. Control variables which measure word meanings i.e., Word Knowledge and I.Q. will have inflated correlations with the M.A.T. Reading because they are measuring some of the same things. Furthermore, since they require Ss to read sentences to do the test, all three tests are measuring the comprehension of sentences. Thus the three control variables have a lot in common with the M.A.T. Reading as well as with each other.

To summarize, the control variables do not appear to be very clean measures of the variables they purport to measure. They have much in common with each other as well as with the Cloze and the M.A.T. Reading comprehension tests.

Measurement of the Variables for Hypothesis 2

Hypothesis 2: Children's skill in making a lexical analysis of the main verb of sentences is positively related to their reading comprehension skill.

Independent Variable: Skill at making a lexical analysis of the main verb of sentences.

This variable will be measured in two ways: first by a sentence completion test, second by looking at Ss' performance on two subsets of paired items on the D.S.R.T.

The Sentence Completion Test consists of the beginning of six sentences up to and including the verb. Ss are asked to complete these sentences with as many words as possible.

The Sentence Completion Test

1. The boy liked ____________________________
2. The girl expected_________________________
3. The boy loved ____________________________
4. The girl remembered_______________________
5. The boy helped____________________________
6. The girl believed__________________________

The verbs in these sentences can take either a transitive construction e.g., The boy liked the girl or a complement construction e.g., The boy liked to go swimming. If it is assumed that complement structures are more complex (See Fodor, Garrett & Bever, 1968) and that Ss with more skill at making a lexical analysis of the verb will produce more complement sentences than Ss with less skill, then the number of complement sentences produced will constitute a measure of skill at making a lexical analysis of the verbs of sentences. The measure employed for this variable is then the number of complement sentences produced in the sentence completion test.

The second measure of Ss' skill at making a lexical analysis of verbs consisted of ten items from the D.S.R.T. Two types of items were used. One type consisted of three paired items with the same sentence structure. One item of each pair had a transitive verb which can take only one deep structure and the other item had a complement verb which can take either a transitive or complement deep structure. Table 9 presents these items.

TABLE 9

Subset of D.S.R.T. Items: Transitive Verbs - Complement Verbs

Transitive (bit)	14.	a. The fat girl bit the thin boy. b. The girl who is thin bit the boy who is fat. c. The boy who is thin was bitten by the girl who is fat.
Complement (loved)	7.	a. The fat girl loved the thin boy. b. The girl who is thin loved the boy who is fat. c. The boy who is thin was loved by the girl who is fat.
Transitive (hit)	5.	a. The girl who the boy hit fell down. b. The boy the girl hit fell down. c. The boy who the girl hit fell down.
Complement (knew)	18.	a. The girl who the boy knew went home. b. The boy the girl knew went home. c. The boy who the girl knew went home.
Transitive (kicked)	11.	a. The tall boy kicked the short girl. b. The boy who is tall kicked the girl who is short. c. The tall boy was kicked by the girl who is short.
Complement (liked)	23.	a. The tall boy was liked by the girl who was short. b. The tall boy liked the short girl. c. The boy who is tall liked the girl who is short.

The second type of item consisted of 2 sets of paired items containing the complement verbs "tell" and "ask," in the same sentence structures. The verb "ask" is more complex than "tell" because it can take more deep structures. (Carol Chomsky, 1968). Table 10 contains these items.

TABLE 10

Subset of D.S.R.T. Items: Tell vs. Ask

Tell #1	21.	a. The boy told the girl when to leave. b. The boy told the girl when he should leave. c. The boy told the girl when she should leave.
Ask #1	4.	a. The girl asked the boy when to leave. b. The girl asked the boy when she should leave. c. The girl asked the boy when he should leave.
Tell #2	10.	a. The boy told the girl what to do. b. The boy told the girl what he should do. c. The boy told the girl what she should do.
Ask #2	25.	a. The boy asked the girl what she should do. b. The boy asked the girl what to do. c. The boy asked the girl what he should do.

If it is assumed that Ss with less skill at making a lexical analysis of verbs will tend to make more errors on items with complex verbs than on items with simpler verbs, then differences in performance on the paired items will constitute a measure of skill at making a lexical analysis of main verbs of sentences.

Dependent Variable: Reading Comprehension Skill.

This variable was measured in the same way as for hypothesis 1.

CHAPTER IV

Description of the Sample and Testing Procedures

Description of the Sample

The original sample for the study consisted of 110 5th grade students in five classrooms in two schools in a suburban community near Boston. When those Ss were eliminated who had missing data and who had word recognition problems, the sample size was reduced to 87. Table 11 presents the basic statistics on this sample.

TABLE 11

Basic Statistics of the Sample (N=87)

Variable	$\bar{X}$	Median	S.D.	Range	Skewness
Age (mos.)	128.034	128.00	4.392	112-140	-0.235
I.Q.	117.011	114.00	12.974	101-150	0.988
Cloze Test (T score)	150.000	150.85	24.712	87-205	-0.250
M.A.T.[a]					
Word Knowledge	4.566 (6.4)	4.3 (6.1)	1.005	2.4-6.5	0.240
Word Discrim.	4.714 (5.5)	4.7 (5.5)	0.886	2.3-5.7	-0.722
Reading	4.984 (5.4)	5.2 (5.8)	1.112	2.2-6.5 (2.7-8.3)	-0.605

[a] The M.A.T. scores are reported in local norms. National norms are presented in parentheses. The expected grade level on national norms is 5-2 since the Ss were tested in the 2nd month of the fifth grade.

The sample is above average in I.Q., above grade level in Word Knowledge, Word Discrimination and Reading according to the national norms.

The sample consisted of 42 males and 45 females. The basic statistics by sex on the relevant variables plus "t" tests for differences between sexes are presented in Table 12.

TABLE 12

Basic Statistics of Sample by Sex

	Males (N=42)		Females (N=45)			
Variable	$\bar{X}$	S.D.	$\bar{X}$	S.D.	t	Prob.
Age (mos.)	128.333	4.812	127.756	3.995	.611	0.543
I.Q.	117.881	14.407	116.200	11.585	.602	0.549
Cloze Test	48.571	10.331	51.111	8.585	-1.233	0.221
M.A.T.						
Word Knowledge	4.588 (6.4)[a]	1.061	4.544 (6.3)[a]	0.961	.201	1.000
Word Discrimination	4.683 (5.5)[a]	0.877	4.742 (5.5)[a]	0.903	.308	1.000
Reading	4.993 (5.4)[a]	1.199	4.976 (5.4)[a]	1.037	.072	1.000

[a] National norms

As can be seen from Table 12 there were no significant sex differences on these variable.

Testing Procedures

Ss were tested in classroom groups whose size ranged from 15 to 32. There were two testing sessions which lasted approximately 45 minutes each. The tests were given in the following order: Sentence Completion, Cloze #1, Cloze #2, Cloze #3 and D.S.R.T. A brief version of the directions for these tests is included here.

1. Sentence Completion Test

   Ss were told to finish the sentence with as many words as they could. There was no time limit on this test.

2. Cloze Tests #1 and #2

   Ss were told to fill in the missing word after going over some samples. There was no time limit set on these tests.

3. Cloze Test #3

   Ss were given the intact paragraph and told to read it silently. They were also told that they would be doing a Cloze test on the paragraph. After 2 minutes they were told to stop and the paragraphs were collected. Then Cloze Test #3 was given out and the students followed the same procedure for Cloze Tests #1 and #2.

4. The D.S.R.T.

   The sample items were completed by the whole class and then discussed to make sure the Ss understood the directions.

Ss were told to do the rest of the items as fast as they could while still being accurate. In addition they were told not to go back and change answers in order to get their first response. After five minutes Ss were told to stop and circle the item that they had just completed. Then they were allowed to finish the test with no time limit.

CHAPTER V

Results and Analysis

Hypothesis 1

Hypothesis 1: Children's skill at recovering the deep structure of sentences is positively related to reading comprehension skill.

This hypothesis was tested by a correlational and regression analysis. First the simple zero order correlations between the D.S.R.T. and the dependent variable, Reading Comprehension, were calculated. Second the relationship between the D.S.R.T. and Reading Comprehension was examined with other variables that could explain the relationship controlled through the use of partial correlations. Third, a regression analysis was performed to determine the relative contribution of the independent variable and control variables in accounting for the dependent variable. Finally the complete analysis was repeated for the male and female subgroups to examine sex differences.

Zero Order Correlations

Hypothesis 1 predicts significant and positive correlations between the D.S.R.T. and Reading Comprehension. Table 13 presents these correlations.

TABLE 13

Correlations between D.S.R.T. and Reading Comprehension (N=87)

Variable	Cloze	M.A.T. Reading
D.S.R.T.	.732**	.476**

$^{**}p < .01$

Clearly the results conform to the prediction.

The correlations between the D.S.R.T. and the Cloze Test are significant and quite large, with more than 50% of the variance accounted for by the D.S.R.T. The relationship between the D.S.R.T. and the M.A.T. Reading is significant but not as great as for the Cloze Test. For the M.A.T. Reading approximately 23% of the variance is accounted for by the D.S.R.T. This lower correlation is to be expected because as has been argued in Chapter IV this traditional type of comprehension test measures much more than reading comprehension. Thus a variable like the D.S.R.T. would not be expected to exhibit as high a degree of relationship to a traditional test as to a more direct measure of comprehension such as the Cloze test. Furthermore, it must be reiterated here that the Cloze test is the measure of the dependent variable for this study for the reasons discussed in Chapter IV. The M.A.T. Reading is only included to show the results on a traditional comprehension test but not as a test of the hypothesis.

When these correlations are attenuated to correct for the unreliability of the tests, the correlations increase to .864 for the Cloze test and to .566 for the M.A.T. Reading. Thus, if the tests were perfectly reliable the D.S.R.T. would account for 74% of the variance in the Cloze test and 32% in the M.A.T. Reading.

In general, reliabilities tend to be so high to begin with that it is difficult to raise them very much. Consequently attenuated correlations aren't usually possible to attain in reality. However, the relatively low reliability of the D.S.R.T., .796, indicates that its reliability can probably be improved through adding more items. Thus it is realistic to assume that correlations approaching the attenuated correlations are attainable. The analysis presented here will employ the unattenuated correlations. Thus the findings will tend to be conservative and may underestimate the true relationship between the D.S.R.T. and reading comprehension.

In summary, the zero order correlations offer strong support for hypothesis 1.

Partial Correlations

The zero order correlations do not tell the whole story. A positive zero order correlation between two variables may be completely explainable by a third variable with which both variables are correlated. Thus the original correlation may be completely spurious and the third variable may be the important one. To examine this possibility, partial correlations were run between the D.S.R.T. and the Cloze test and the M.A.T. Reading with each of the three control variables partialled out. Table 14 presents the intercorrelations of all the variables used in testing hypothesis 1. Table 15 presents the partial correlations.

TABLE 14

Intercorrelations of all Variables Used in Testing Hypothesis 1* (N-87)

Variable	1	2	3	4	5	6
1. D.S.R.T.	1.000	.541	.455	.433	.476	.732
2. Word Knowledge		1.000	.674	.761	.740	.621
3. I.Q.			1.000	.585	.587	.484
4. Word Discrimination				1.000	.663	.542
5. M.A.T. Reading					1.000	.528
6. Cloze Test						1.000

*All correlations are significant at the .01 level.

TABLE 15

Partial Correlations: D.S.R.T. and Reading Comprehension with I.Q., Word Discrimination (W.D.), Word Knowledge (W.K.) Controlled

Variable	0 order r	Partial Correlations		
		I.Q.	W.D.	W.K.
D.S.R.T. vs.				
Cloze Test	.732**	.658**	.657**	.597**
M.A.T. Reading	.476**	.280**	.290**	.133

$^{**}p < .01$

The partial correlations show that the control variables do not explain much of the relationship between the D.S.R.T. and the Cloze test of comprehension. Very little of the relationship between the D.S.R.T. and the Cloze test can be accounted for by I.Q., Word Discrimination, and Word Knowledge. This data provides further strong confirmation for hypothesis 1.

The partial correlations for the M.A.T. Reading show that the control variables explain much of the relationship between the D.S.R.T. and the M.A.T. Reading. This is to be expected since these variables have much in common with the M.A.T. Reading test.

Regression Analysis

It is important to inquire into the relative importance of the D.S.R.T. and the control variables in explaining reading comprehension. A regression analysis can be employed to conduct this inquiry. A short discussion of this type of analysis follows.

A regression analysis attempts to find the optimal weighted combination of a number of independent variables in predicting a dependent variable. The degree of predictability that the independent variables provide can be looked at in terms of the amount of variance in the dependent variable that is accounted for by the independent variables. Thus the total variance in regression analysis can be conceived of as consisting of two parts -- the explained variance and the unexplained variance. The explained variance can be further broken down into two parts -- the unique variance and the common variance. These four components are described below.

Explained Variance This is the amount of variance that is accounted for by all the independent variables. It is the sum of the unique and common variance.

The Unique Variance This is the amount of variance that a given variable does not share with the other variables. It is specific to a particular variable.

The Common Variance This is the variance that is shared by all the independent variables. It is what is common to all of them.

Unexplained Variance This is the amount of variance that cannot be predicted by the variables. It includes error variance plus variance that is due to other variables not measured in the regression analysis.

These components of variance add up to 100% which is the total variance. The variance explained is simply the multiple R^2 with all predictor variables entered into the regression equation. The unexplained variance is the multiple R^2 subtracted from 100.

The relative importance of a variable can be looked at in two ways. First the unique variance attributable to that variable can be looked at. This figure provides an estimate of the unique importance of a variable i.e., how much variance it can explain that cannot be explained by the other variables. The unique variance is found by entering all the variables except a specified variable into the regression first. Then the increase in R^2 found when the specified variable is entered into the regression gives the unique variance of that variable. A second estimate of the relative importance of a variable can be obtained by adding to the unique variance of that variable a share of the common variance that is in proportion to its zero order correlation with the dependent variable. This estimate is obtained by multiplying the beta weight of each variable by its zero order correlation with the dependent variable.

Table 16 presents the results of the regression analysis with the Cloze test as the dependent variable.

TABLE 16

Regression Analysis: D.S.R.T., Word Knowledge, Word Discrimination, I.Q. with Cloze Test (N=87)

Variable	(1) 0-order r	(2) Beta Weight	% Unique + proportion Common (1) x (2)	% Unique Variance	Mult R	% Variance Explained
D.S.R.T.	.732	.5547	40.6	21.3	.785	61.6
Word Knowledge	.621	.2118	13.2	01.4		
Word Discrimination	.542	.1351	07.3	00.8		
I.Q.	.484	.0097	00.5	00.0		
			Total Unique	23.5	% Unexplained Variance	38.4
			Common Variance	38.1		

Table 16 shows that the D.S.R.T. is the only variable in the regression that has any unique variance worth discussing. And the amount of unique variance (21.3%) is very substantial. The analysis of the unique plus common variance shows that the D.S.R.T. accounts for 40.6% of the explained variance which is quite substantial. This percentage is three times as great as the next highest variable which contributes 13.2% and over 5 times as great as the next highest variable Word Discrimination (7.3%). I.Q. makes almost no contribution at all.

The results of the regression analysis demonstrate that the D.S.R.T. is, by a substantial amount, the most important factor in reading comprehension as measured by the Cloze test when compared to the other variables in the analysis.

As expected when the M.A.T. Reading is the dependent variable the situation changes as Table 17 shows.

TABLE 17

Regression Analysis: D.S.R.T., Word Discrimination, Word Knowledge, I.Q. with M.A.T. Reading (N=87)

Variable	(1) 0-order r	(2) Beta Weight	% Unique + proportion Common (1) x (2)	% Unique Variance	Mult R	% Variance Explained
D.S.R.T.	.476	.0851	04.1	0.5	.765	58.6
Word Knowledge	.740	.4525	33.5	6.4		
Word Discrimination	.663	.2121	14.1	1.9		
I.Q.	.587	.1137	06.7	0.7		
			Total Unique	9.5	% Unexplained Variance 41.4	
			Common Variance	49.1		

With the M.A.T. Reading as the dependent variable, Word Knowledge accounts for the highest amount of unique variance (6.4%). However this amount is quite small. The other variables account for hardly any unique variance. In fact all the variables taken together only account for 9.5% of the unique variance, leaving a common variance of 49.1%. This means that almost 5/6ths of the variance explained is common variance i.e., variance common to all the tests. This adds support to the suggestion in Chapter IV that the Word Knowledge, Word Meaning and I.Q. tests are measuring the same thing. The dependent variable is so global as are three of the independent variables that it is not surprising to find a lot of common variance. The unique plus common variance distribution shows that Word Knowledge is (33.5%), slightly more than twice as important as Word Discrimination (14%), which in turn is nearly 3 times as important as either I.Q. or the D.S.R.T. Thus Word Knowledge seems to be the most important factor in the M.A.T. Reading.

To summarize, the evidence from the zero order and partial correlations overwhelmingly confirms hypothesis 1. In addition the regression analysis suggests that the D.S.R.T. is a very important factor in comprehension, a factor not accounted for by other variables such as Word Knowledge skill, Word Recognition skill, and I.Q. The regression analysis for the total group provides strong evidence that the recovery of deep structure is an important aspect of reading comprehension.

Sex Subgroup Analysis

In order to see if hypothesis 1 had differential effects, the total sample was broken into a male and a female subgroup and the analysis was repeated for each subgroup. As can be seen from Table 18 the only sex differences in mean scores on the variables was on the D.S.R.T. On all other variables there are no differences.

TABLE 18

Comparison of Means for Males and Females

Variable	Males (N=42)		Females (N=45)			Prob.
	$\bar{X}$	S.D.	$\bar{X}$	S.D.	t	(two-tailed)
D.S.R.T.	17.452	4.278	19.622	4.000	2.44	.01
Word Knowledge	4.588	1.061	4.544	.961	0.201	N.S.
I.Q.	117.881	14.407	116.200	11.585	0.600	N.S.
Word Discrimination	4.683	0.877	4.742	.903	0.300	N.S.
M.A.T. Reading	4.922	1.199	4.976	1.037	0.070	N.S.
Cloze Test	48.571	10.331	51.111	8.858	1.233	N.S.

Note. - N.S. = nonsignificant

Zero Order Correlations:

Table 19 presents the zero order correlations between the D.S.R.T. and Reading Comprehension by sex:

TABLE 19

Correlations between D.S.R.T. and Comprehension by Sex

Variable	Cloze		M.A.T. Reading	
	Male	Female	Male	Female
D.S.R.T.	.768**	.682**	.682**	.279

** $p < .01$.

These correlations support hypothesis 1 for both males and females because of the significant and substantial correlation between the D.S.R.T. and the Cloze Test. 58.9% of the variance is accounted for by the D.S.R.T. for males and 46.5% for females. The difference between the correlations for males and females is not statistically significant.

The attenuated correlations are .864 for males and .805 for females. These correlations indicate that if the tests were perfectly reliable 74% of the variance in the Cloze test could be explained for males and 64% for females. On the M.A.T. Reading the D.S.R.T. is significantly correlated for males but not for females.

Partial Correlations The partial correlations by sex are presented in Table 20.

TABLE 20

Partial Correlations: D.S.R.T. and Reading Comprehension with I.Q., Word Discrimination (W.D.), Word Knowledge (W.K.) controlled. Analysis by sex.

Variable	0-order r		Partial Correlations					
D.S.R.T. vs.			I.Q.		W.D.		W.K.	
	M	F	M	F	M	F	M	F
Cloze	.768**	.682**	.666**	.569**	.569**	.662**	.510**	.573**
M.A.T. Rdg.	.682**	.279	.438**	.020	.385*	.186	.213	-.108

*p $<$ .05 **p $<$ .01

The partial correlations between D.S.R.T. and the Cloze test of reading comprehension with I.Q., Word Discrimination, and Word Knowledge separately controlled show significant correlations for both males and females. Taken individually the control variables do not explain the relationship between the D.S.R.T. and the Cloze test of comprehension. Thus for both male and female subgroups hypothesis 1 is confirmed.

On the M.A.T. Reading the partial correlations are not significant for females but are significant for males with I.Q. and Word Discrimination controlled. With Word Knowledge controlled the partial correlation becomes non significant for males also.

Regression Analysis The intercorrelations of all the variables used in testing hypothesis 1 are presented by sex in Table 21 (p.86). These correlations show the same general pattern as for the total group. The correlations for the males appear to be generally higher than for the females.

The results of the regression analysis with the Cloze comprehension test as the dependent variable are presented in Table 22 (p.87).

Only the D.S.R.T. contributes any substantial degree of unique variance. The unique variance contributed by the females is nearly 10% more than for the males. As for the common plus unique variance the D.S.R.T. accounts for nearly 2/3rds of the total variance accounted for in both male and female subgroups. This is nearly twice as much as the next highest variable. Next to the D.S.R.T., Word Knowledge seems to be more important for females and Word Discrimination for males. The results of this analysis by sex is more or less the same as for the total group. For both males and females the D.S.R.T. appears to be a very important unique aspect of reading comprehension as measured by the Cloze test.

The results of the regression analysis for sex subgroups with the M.A.T. Reading as the dependent variable are presented in Table 23.

TABLE 21

Intercorrelations of All Variables Used in Testing Hypothesis 1, by Sex:
Males (N=42) Females (N=45)

Variable	1		2		3		4		5		6	
	M	F	M	F	M	F	M	F	M	F	M	F
1. D.S.R.T.	1.000	1.000	.724	.392	.547	.419	.662	.224	.682	.279	.768	.682
2. Word Knowledge			1.000	1.000	.692	.655	.835	.693	.706	.781	.683	.571
3. I.Q.					1.000	1.000	.624	.559	.632	.529	.516	.473
4. Word Discrimination							1.000	1.000	.735	.593	.702	.380
5. M.A.T. Reading									1.000	1.000	.606	.440
6. Cloze Reading												

*Correlations over .288 are significant at the .05 level. Correlations over .338 are significant at the .01 level.

TABLE 22

Regression Analysis by Sex: D.S.R.T., Word Knowledge,

Word Discrimination, I.Q. with Cloze. Males (N=42) Females (N=45)

Variable	(1) O-order r		(2) Beta Weight		% Unique + proportion Common (1)x(2)		% Unique Variance		Mult R		% Variance Explained	
	M	F	M	F	M	F	M	F	M	F	M	F
D.S.R.T.	.768	.682	.5347	.5393	41.1	36.8	13.2	22.9	.810	.758	65.7	57.5
Word Knowledge	.683	.571	.0150	.3385	01.0	19.3	00.0	04.6				
Word Discrim.	.702	.380	.3317	.0151	19.9	00.6	03.2	00.0				
I.Q.	.516	.473	.0062	.0169	00.3	00.3	00.0	00.0				
					Total Unique		16.4	27.5	% Unexplained Variance		34.3	42.5
					Common Variance		49.3	30.0				

TABLE 23

Regression Analysis by Sex: D.S.R.T., Word Knowledge, Word Discrimination, I.Q., with M.A.T. Reading

Variable	(1) O-order r		(2) Beta Weight		%Unique + proportion Common (1)x(2)		% Unique Variance		Mult R		% Variance Explained	
	M	F	M	F	M	F	M	F	M	F	M	F
D.S.R.T.	.682	.279	.2908	-.0309	19.8	00.0	03.9	00.0	.798	.785	63.6	61.6
Word Knowledge	.706	.781	.0178	.7151	01.3	55.8	00.0	20.3				
Word Discrim.	.735	.593	.3943	.0925	29.0	05.5	04.5	00.4				
I.Q.	.632	.529	.2145	.0215	13.6	01.1	02.3	00.0				
					Total Unique		10.7	20.7	% Unexplained Variance		36.4	38.4
					Common Variance		52.9	39.9				

For males very little unique variance is accounted for by any of the variables. When unique and common variance are combined, the D.S.R.T. and Word Discrimination are the two most important variables. For females Word Knowledge is the most important variable accounting for 20.3% of the unique and 55.8% of the unique plus common variance. The other variables don't contribute very much if at all. For males the D.S.R.T. and Word Discrimination are important and for females Word Knowledge is the only important variable. The D.S.R.T., as a factor in reading comprehension, as measured by the M.A.T. Reading, is important for males and of little importance for females.

Hypothesis 1: Conclusions

The results of the analysis of hypothesis 1 yield the following conclusions.

1. Hypothesis 1 is supported by the results. The results of the analysis show that Ss' ability to recover deep structure is related to reading comprehension.

2. Recovering the deep structure is an important aspect of reading comprehension. In fact Ss' skill at recovering the deep structure of sentences is a much more important aspect of reading comprehension skill, as measured by a Cloze test, than I.Q., word knowledge and word recognition skill. Furthermore only Ss' ability to recover deep structure makes any substantial unique contribution to reading comprehension skill.

3. Word knowledge is a more important aspect of reading comprehension, as measured by a traditional comprehension test, than word recognition, I.Q. and skill at recovering deep structure. However, none of these factors makes very much unique contribution to reading comprehension as measured by a traditional comprehension test.

4. Females are superior to males in skill at recovering the deep structure of sentences.

5. There are no sex differences in the importance of skill at recovering deep structure as an aspect of reading comprehension when reading comprehension is measured by the Cloze test.

6. For both male and female subgroups, Ss' skill at recovering the deep structure of sentences is a much more important aspect of reading comprehension, as measured by the Cloze test, than I.Q., word knowledge and word recognition skill.

7. For males, word recognition skill and skill at recovering the deep structure are both more important aspects of reading comprehension, as measured by the M.A.T. Reading test, than word knowledge and I.Q. However none of these variables makes very much unique contribution to reading comprehension skill.

8. For females word knowledge is a much more important aspect of reading comprehension, as measured by the M.A.T. Reading test, than I.Q., word recognition skill and skill at recovering the deep structure. Furthermore for females word knowledge makes the only substantial unique contribution

to reading comprehension.

9. The findings for the Cloze criterion appear to hold up under replication. The testing took place in two different schools, and the testing in the second school constitutes a replication of the testing in the first school. The analysis was repeated for each school separately. The results as reported in Appendix D for both schools support hypothesis 1.

Hypothesis 1: Discussion

The differences in the results for hypothesis 1 between the Cloze test criterion and the M.A.T. Reading test criterion deserves further comment. For the total group Ss' skill in recovering deep structure appears to be an important aspect of reading comprehension when reading comprehension is measured by the Cloze test. When reading comprehension is measured by the M.A.T. Reading test, Ss' ability to recover deep structure, although related to reading comprehension, does not appear to be an important aspect of it as compared to I.Q., word discrimination and word knowledge skill. The two ways of measuring comprehension then produce different results.

The difference in results can be explained by problems in the measurement of variables. As discussed in Chapter III, the I.Q. test, the Word Discrimination test and the M.A.T. Reading test all seem to be measuring the same thing, thus they seem to have a great deal of common variance. This is exactly what Table 17 shows -- a great deal of common variance and very little unique variance. This shared variance appears to be due in part to the measurement procedure, i.e. standardized multiple choice tests, rather than to the intrinsic relationship of the variables.

Furthermore the fact that the M.A.T. Reading test, Word Knowledge test and Word Discrimination test are subtests of the same standardized test guarantees that they will be correlated with each other. All this plus the global nature of the variables insures a great deal of common variance. It is thus difficult for a relatively clean cut measure like the D.S.R.T. to account for any unique variance in the M.A.T. Reading test. The results gathered from the M.A.T. Reading test criterion are difficult to interpret because of these measurement problems. Therefore, the use of the Cloze test criterion constitutes the only valid test of hypothesis 1.

The analysis of the sex subgroups also deserves comment. The only differences between the mean scores for males and females is on the D.S.R.T. where females are superior to the males. On all other variables there is no statistical significance between males and females. The superiority of the females over the males in ability to recover the deep structure of sentences may be due to the females' greater skill or to some extraneous aspect of the test or testing situation. This difference, if it were upheld under replication, would deserve further study.

Another sex difference is in the relationship between the D.S.R.T. and the M.A.T. Reading test. For males the D.S.R.T. is related to the M.A.T. Reading test, but for females there is no relationship. This difference is not easy to explain.

At first glance then it appears that the ability to recover deep structure is an important aspect of reading comprehension for males but not for females. This might be explained through the following speculation. The ability to recover deep structure is a lower level skill than some of the higher level skills that traditional reading tests claim they are measuring. Since males are usually lower than females in reading ability at the beginning stages of reading, it is possible that the males are at a lower stage of reading development than the females. And at this lower stage, where the males are, the ability to recover the deep structure is important, but at the higher stages, where the girls are, it is not important. This would explain the significant correlation for males and the lack of significant correlation for females between the D.S.R.T. and the M.A.T. Reading test.

This speculation, however, is not borne out by the data. The notion that the females are on a higher level and that the males are on a lower level does not seem to hold for the sample studied here. First, there are no mean differences between the males and females on all the variables except the D.S.R.T. Thus the data does not indicate that females generally are at a higher level than males. Second, there are no sex differences in the strength of the relationship between the D.S.R.T.

and the Cloze test. Third, there is no theoretical reason to expect sex differences.

It appears that the sex differences with the M.A.T. Reading criterion is due to a pecularity of this test or of the sample studied. What pecularity of the sample or the M.A.T. Reading test is at work here remains an open question. It is very possible that the difference will disappear under replication. If this difference is maintained under replication, however, it should be investigated further.

Hypothesis 2

Hypothesis 2: Children's skill in making a lexical analysis of the main verb of a sentence is positively related to their reading comprehension skill.

This hypothesis can be tested in two ways. The first way is through correlations between the sentence completion test and reading comprehension. The second way is by examining through t tests the relative performance of good and poor readers on the subset of items on the D.S.R.T. as discussed in Chapter III.

Sentence Completion Test

The first test of hypothesis 2 can be made by correlating the number of complement sentences produced by Ss on the Sentence Completion test with reading comprehension skill. Hypothesis 2 predicts a positive and significant correlation.

These correlations are presented in Table 24.

TABLE 24

Correlations: Sentence Completion Test vs. Reading Comprehension (N=87)

	Cloze	M.A.T. Reading	$\bar{X}$	S.D.
Sentence Completion (Number of Complement Sentences)	.270*	.191	3.437	1.717

* $p < .05$

The results show confirmation for hypothesis 2. The correlation between the Sentence Completion test and the Cloze test is statistically significant but low. On the M.A.T. Reading it is non significant.

A repeat of the analysis for male and female subgroups is presented in Table 25.

TABLE 25

Correlations: Sentence Completion vs. Reading Comprehension by Sex

Sentence Completion	Cloze	M.A.T. Reading	$\bar{X}$	S.D.	N
Males	.389*	.325*	3.048	1.847	42
Females	.069	.031	3.800	1.517	45

*$p < .05$

The hypothesis is confirmed for males but not for females. The significant correlation for the total group is due to the male subgroup.

Item Subsets from the D.S.R.T.

The second test of hypothesis 2 involves looking at the performance of good and poor readers on the subset of the D.S.R.T. items containing lexically simple and complex verbs. This involves dividing the group into good readers, i.e. upper 40% on the Cloze test, and poor readers, i.e. lower 40% on the Cloze test. Next a comparison of performance for each of these groups on the items containing the simple verbs, i.e. transitive and tell verbs versus items containing more complex verbs, i.e. complement and ask verbs, is made through the use of t tests. The expected performance is as follows:

	Simple Verbs		Complex Verbs	t-test
Good Readers (Upper 40% Cloze test)	$\bar{X}$	>	$\bar{X}$	
Poor Readers (Lower 40% Cloze test)	$\bar{X}$	$\cong$	$\bar{X}$	t-test
Total Group	$\bar{X}$	>	$\bar{X}$	t-test

These predictions are based on the following line of reasoning:

Total Group Simple verbs will be easier than complex verbs according to the lexical analysis strategy for recovery of deep structure as proposed by Fodor, Garrett and Bever 1968 and discussed in Chapter II.

Good Readers These readers tend to employ the lexical analysis strategy and thus will find simple verbs easier than complex verbs.

Poor Readers The readers because of less skill at employing the lexical analysis strategy will find the complex verbs and simple verbs equally difficult.

Table 26 presents the results of this analysis.

TABLE 26

t-tests
Simple vs. Complex Verbs by Reading Comprehension (Upper 40%/Lower 40%)

	Simple Verbs (Transitive)			Complex Verbs (Complement)					
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	Diff	t	Prob. (one-tailed)
High Cloze (Upper 40%)	2.611	.494	36	2.444	.558	36	.167	1.435	<.10>.05
Low Cloze (Lower 40%)	1.800	.933	35	1.771	.877	35	.029	0.128	>.10
Total Group	2.218	.827	87	2.069	.818	87	.149	1.312	<.10>.05

	Simple Verbs (Tell)			Complex Verbs (Ask)					
	$\bar{X}$	S.D.	N	$\bar{X}$	S.D.	N	Diff	t	Prob. (one-tailed)
High Cloze (Upper 40%)	1.917	.280	36	1.778	.485	36	.139	1.405	<.10>.05
Low Cloze (Lower 40%)	1.371	.690	35	1.443	.550	35	.229	1.675	≅ .06
Total Group	1.644	.590	87	1.488	.605	87	.195	2.557	<.02

The predictions for the total group are supported in the case of the Tell-Ask verbs where the difference is in the predicted direction and this difference is significant. In the case of the Transitive and Complement verbs the difference is in the right direction but is not significant. In this case the prediction is not upheld.

The predictions for the good readers are not supported. Although the transitive verbs are easier than the complement verbs and the Tell verbs are easier than the Ask verbs, the differences are not statistically significant. These differences do however approach the .05 level.

The predictions for the poor readers are generally supported in that there are no statistically significant differences between simple and complex verbs. However for the poor readers the Tell-Ask difference is very close ($p \cong .06$) to being significant. This of course is contrary to the predictions of equal difficulty.

In neither case, i.e. good readers or poor readers, are the predictions completely upheld by the data.

In the case of the transitive and complement verbs the results are in the right direction but not significant. In the case of the Tell and Ask verbs the data does not support the hypotheses, so that overall the analysis of simple and complex verbs does not really support hypothesis 2.

Hypothesis 2: Conclusions

The results of the analysis offer only slim support for hypothesis

2. Skill at making a lexical analysis of the main verb of the sentence as measured in this study appears to be only slightly, if at all, related to reading comprehension. The results of the analysis show:

1. Skill at making a lexical analysis of the main verb of a sentence, as measured by the Sentence Completion test, shows a small but statistically significant relationship to reading comprehension, as measured by the Cloze test, and no relationship when comprehension is measured by the M.A.T. Reading test.

2. For males skill at making a lexical analysis of the main verb of sentences, as measured by the Sentence Completion test, is moderately related to reading comprehension, as measured by the Cloze test or the M.A.T. Reading.

3. For females skill at making a lexical analysis of the verb, as measured by the Sentence Completion Test, shows no relationship to reading comprehension, as measured by the Cloze test or the M.A.T. Reading.

4. Skill at making a lexical analysis of the main verb of a sentence, as measured by the subset of paired items on the D.S.R.T., is not related to reading comprehension, as measured by the Cloze test.

Hypothesis 2: Discussion

The results of the analysis of hypothesis 2 can only be described as mixed. The results of the first test of hypothesis 2, the Sentence Completion test, support the hypothesis but in a weak way. The second test of the hypothesis, the analysis of simple and complex verbs, shows essentially negative results.

The second test of hypothesis 2 deserves further discussion. Its negative results can be explained in two ways. First, the hypothesis could be wrong. If this is the case, then there is obviously no reason to expect positive results. A second explanation could be that the hypothesis is correct but the instruments used to measure the predicted effects are not sensitive enough to show the effects. The second possibility seems plausible for several reasons. First the demonstration of the lexical analysis strategy as reported by Fodor et al. (1968) was a weak one. In other words the lexical analysis strategy is difficult to demonstrate and requires sensitive instruments and carefully controlled experimentation. Fodor used a reaction time criterion while the present study used an error criterion. In the study reported here the lexical analysis strategy was demonstrated for the total group only for the tell-ask verbs but not for the transitive-complement verbs. Thus the failure to replicate the Fodor et al. results, i.e. that complex verbs are more difficult to comprehend than simple verbs, leads one to believe that the measuring instrument used, which was different from Fodor's, was not sensitive enough to pick up the differences in processing. In short, it is possible that the lack of positive results for hypothesis 2 could be explained by the lack of sensitivity of the instruments measuring Ss' use of the lexical analysis strategy.

CHAPTER VI

Implications

The recovery of deep structure has been identified in this paper as an important aspect of the reading comprehension process. The unambiguous nature of this concept renders it an excellent point of departure for further research in reading comprehension. This should in turn lead to new instructional methods.

Research Implications

Research into Ss' ability to recover deep structure should take the following direction.

1. The present study should be replicated to determine if the results can be generalized beyond the sample studied.
2. When this study is replicated, other measures of Ss' ability to recover deep structure should be developed to make certain that the results reported here are not due to pecularities of the D.S.R.T. rather than to skill in recovering deep structure.
3. The developmental aspects of Ss' ability to recover deep structure should be investigated to determine how this skill is acquired.
4. Research should be undertaken to investigate the strategies that Ss use in recovering the deep structure. This research

could follow three lines. First, other strategies, like the lexical analysis strategy, need to be developed. These strategies will probably come from the work in psycholinguistics. Second, research should be undertaken to distinguish which of these strategies are specific to reading and which are part of language comprehension in general. Finally, research should be undertaken to investigate the pertinence of these strategies to differences in reading comprehension skills. This research should take the form of studies aimed at determining whether or not differences in Ss' skill in utilizing various strategies differentiate good and poor readers.

5. Instructional research should be undertaken to determine if the strategies that Ss use in recovering the deep structure of sentences are amenable to instruction. And if improvement in these strategies will improve reading comprehension.

Instructional Implications

Increased knowledge of the strategies Ss use in recovering deep structure should lead to the development of diagnostic tests to identify children who are having difficulty in recovering deep structure and to pinpoint the deficient aspect of the deep structure recovery process. Instructional procedures and materials could then be developed to remedy the particular problems identified by the diagnostic tests. It is impossible to name these instructional procedures now since they will be determined by the types of strategies -- at present unknown -- that Ss use when recovering deep structure.

In this author's opinion it would be ill advised at present to develop instructional procedures and materials before more is known about these strategies. To do so would simply be to repeat the mistakes made in the past when research was applied to instruction (see Chapter I). It would be incorrect to assume that Ss' skill in the recovery of deep structure could be improved simply by providing them practice in this skill through materials like the D.S.R.T. Such an approach would be a typical example of the nontheoretical thinking which currently allows testing procedures to be used as prototypes for teaching procedures.

This approach is wrong for two reasons. First there is no particular evidence that the measurement of a skill is a good model for instruction. Second this approach provides no way of helping students who can not do the task.

It seems to this author the development of instructional procedures must wait until more is known about the actual strategies, not just the general skill, used in recovering deep structure. In this way instruction could be much more specific and directed at particular student deficiencies. Thus increased knowledge of the strategies that Ss use in recovering deep structure could provide the key to instruction in this important skill.

In summary, the findings of this study could lead in two directions.

First they could lead to more research into children's skill at recovering deep structure. Second it could lead to the development of tests to identify students with deficiencies in this skill and to instructional procedures and materials to remedy these deficiencies.

APPENDIX A

Deep Structure Recovery Test (D.S.R.T.)

Name__________________________

Directions: Circle the sentence that has a different meaning.

Sample 1.
a. The boy hit the girl.
b. The girl was hit by the boy.
c. The boy was hit by the girl.

Sample 2.
a. The fat girl looked at the short boy.
b. The short boy looked at the fat girl.
c. The fat girl was looked at by the short boy.

@ Copyright 1969 Herbert D. Simons

Name____________________________

1. a. The boy gave a book to the girl.

   b. The book was given the girl by the boy.

   c. The book was given to the boy by the girl.

2. a. What the boy would like is for the girl to leave.

   b. For the boy to leave is what the girl would like.

   c. What the girl would like is for the boy to leave.

3. a. He painted the red house.

   b. He painted the house red.

   c. He painted the house that was red.

4. a. The girl asked the boy when to leave.

   b. The girl asked the boy when she should leave.

   c. The girl asked the boy when he should leave.

5. a. The girl who the boy hit fell down.

   b. The boy the girl hit fell down.

   c. The boy who the girl hit fell down.

Name________________________

6. a. The boy saw the paper that was folded.
   b. The boy saw the folded paper.
   c. The boy saw the paper folded.

7. a. The fat girl loved the thin boy.
   b. The girl who is thin loved the boy who is fat.
   c. The boy who is thin was loved by the girl who is fat.

8. a. The boy saw his broken toy.
   b. The boy saw his toy that was broken.
   c. The boy saw his toy broken.

9. a. That the girl was right was believed by the boy.
   b. What the girl believed was that the boy was right.
   c. The girl believed the boy to be right.

10. a. The boy told the girl what to do.
    b. The boy told the girl what he should do.
    c. The boy told the girl what she should do.

Name______________________________

11. a. The tall boy kicked the short girl.
    b. The boy who is tall kicked the girl who is short.
    c. The tall boy was kicked by the girl who is short.

12. a. That the girl would go into the house was expected by the boy.
    b. What the girl expected was that the boy would go into the house.
    c. The girl expected the boy to go into the house.

13. a. The boy begged the girl to tell the truth.
    b. What the boy begged the girl to do was tell the truth.
    c. To tell the truth was what the girl begged the boy to do.

14. a. The fat girl bit the thin boy.
    b. The girl who is thin bit the boy who is fat.
    c. The boy who is thin was bitten by the girl who is fat.

15. a. The girl taught the boy to use a pencil.
    b. What the boy taught the girl was to use a pencil.
    c. To use a pencil was what the girl taught the boy.

Name________________________

16. a. The wagon is easy for us to push.

b. We push the wagon easily.

c. The wagon pushes us easily.

17. a. The girl ordered the boy to sit down.

b. What the girl ordered the boy to do was sit down.

c. To sit down was what the boy ordered the girl to do.

18. a. The girl who the boy knew went home.

b. The boy the girl knew went home.

c. The boy who the girl knew went home.

19. a. The girl sees easily.

b. The girl is easy to see.

c. It is easy to see the girl.

20. a. The man found his car stolen.

b. The man found that his car was stolen.

c. The man found his stolen car.

Name____________________________

21. a. The boy told the girl when to leave.
    b. The boy told the girl when he should leave.
    c. The boy told the girl when she should leave.

22. a. The boy sees that the girl is tall.
    b. That the girl is tall is seen by the boy.
    c. What the girl sees is that the boy is tall.

23. a. The tall boy was liked by the girl who was short.
    b. The tall boy liked the short girl.
    c. The boy who is tall liked the girl who is short.

24. a. What the girl wants is for the boy to find the ball.
    b. The boy wants the girl to find the ball.
    c. For the boy to find the ball is what the girl wants.

25. a. The boy asked the girl what she should do.
    b. The boy asked the girl what to do.
    c. The boy asked the girl what he should do.

APPENDIX B

Item Statistics: D.S.R.T. (N=103)

Item Number	Item Difficulty (Proportion correct)	Item Validity (Item vs. Total score -Biserial r)
1	.845	.510
2	.709	.654
3	.282	.556
4	.563	.387
5	.495	.358
6	.485	.648
7	.854	.409
8	.544	.683
9	.738	.711
10	.786	.549
11	.874	.767
12	.709	.573
13	.922	.571
14	.845	.568
15	.903	.743
16	.903	.703
17	.961	.601
18	.631	.436
19	.825	.486
20	.767	.646
21	.825	.700
22	.825	.842
23	.573	.575
24	.874	.491
25	.515	.632

APPENDIX C

Cloze Tests of Comprehension

Cloze Test #1.

A nobleman and a _________ met in a tavern. _________ their lunch they ordered _________. When it was brought, _________ nobleman took a spoonful, _________ the soup was so _________ that he burned his _________ and tears came to _________ eyes. The merchant asked _________ he was weeping. The _________ was ashamed to admit _________ had burned his _________ answered, "Sir, I once _________ a brother who committed _________ great crime, for which _________ was hanged. I was _________ of his death, and _________ made me weep." The _________ believed this story and _________ to eat his soup. _________ too burned his mouth, _________ that he had tears _________ his eyes. The nobleman _________ it and asked the _________ "Sir, why do you _________?" The merchant, who now _________ the nobleman had deceived _________, answered "My lord, I _________ weeping because you were _________ hanged together with your _________.

Cloze Test #2.

Most ants __________ hard workers and often __________ from six o'clock in __________ morning until ten o'clock __________ night. The work is __________ among the worker ants __________ that each one has __________ certain amount to do. __________ do not know how __________ decide what each one __________ to do, for they __________ not talk. Some people __________ ants follow each other __________ their sense of smell. __________ often live to be __________ year old, and some __________ been known to live __________ or seven years. One __________ they get their food __________ from plant lice, which __________might call their cows. __________ ants milk these "cows" __________ tapping the lice gently __________ a drop of honey __________ out. Then they eat __________ honey. Ants take very __________ care of these plant __________ and often they build __________ covering over them so __________ they will be protected __________ the rain.

Cloze Test #3.

We __________ all ready for the __________ with our packs strapped __________ securely. We followed the __________ nearly all day, except __________ we stopped to lunch __________ the side of a __________ cliff. By sundown we __________ anxiously looking for a __________ to camp for the __________. We gathered a huge __________ of eucalyptus branches to __________ ready for our fire. __________ supper we rolled up __________ our blankets near the __________ and we were soon __________. We were awakened by __________ queer sounds which my __________recognized as howls of __________ jaguar. We hastily heaped __________ branches on the fire __________ was burning low. We __________ our ukeleles and sang. __________ means of much light __________ noise we kept all _______ away except an inquisitive __________ and several rabbits and __________. We fell asleep toward __________ and we were not __________ again during our stay __________ the canyon.

APPENDIX D

Regression Analyses for each School

School #1

Regression Analysis: D.S.R.T., Word Knowledge
Word Discrimination, I.Q. with Cloze Test (N=39)

Variable	(1) 0-order r	(2) Beta Weight	(1) x (2) % Unique + proportion common variance	% Unique Variance	Mult R	% Variance Explained
D.S.R.T.	.754	0.6237	47.0	26.5	.791	62.6
Word Knowledge	.584	0.3308	19.3	03.3		
Word Discrimination	.463	-0.0786	-3.6^{a}	00.0		
I.Q.	.519	-0.0019	-0.1^{a}	00.2		
			% Total Unique Variance	30.0	% Unexplained Variance 37.4	
			% Common Variance	32.6		

[a] Word Discrimination and I.Q. act as suppressor variables thus decreasing the multiple R and consequently the amount of variance explained.

School #2

Regression Analysis: D.S.R.T., Word Knowledge Word Discrimination, I.Q. with Cloze Test (N=48)

Variable	(1) 0-order r	(2) Beta Weight	(1) x (2) % Unique + proportion common variance	% Unique Variance	Mult R	% Variance Explained
D.S.R.T.	.698	.4907	34.3	16.3	.797	63.5
Word Knowledge	.658	.1473	09.7	00.6		
Word Discrimination	.614	.3106	19.1	04.4		
I.Q.	.473	.0088	00.4	00.0		
			%Total Unique Variance	30.0	% Unexplained Variance	37.4
			%Common Variance	32.6		

REFERENCES

Anderson, C.C. A factorial analysis of reading. British Journal of Educational Psychology, 1949, 19, 220-221.

Barrett, T.C. Taxonomy of cognitive and affective dimensions of reading comprehension. Discussed by Clymer, T. in "What is reading?": some current concepts. In Helen M. Robinson (Ed.), Innovational change in reading instruction. Sixty seventh yearbook: National Society for Study in Education, 1968, University of Chicago Press. 1-30.

Bever, T.G., Mehler, J., & Carey, P. What we look at when we read. Perception and Psychophysics, 1967, 2, 213-218.

Bleismer, E.P. Reading abilities of bright and dull children of comparable mental ages. Journal of Educational Psychology, 1954, 45, 321-331.

Blumenthal, A.L. Prompted recall of sentences. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 203-206.

Blumenthal, A.L., & Boakes, R. Prompted recall of sentences. Journal of Learning and Verbal Behavior, 1967, 6, 674-675.

Bormuth, J.R. (Ed.) Readability in 1968. National Council of Teachers of English. 1968.

Cafone, H.C. Individual differences in the reading processes of ninth grade retarded readers. Unpublished doctoral dissertation, University of Arizona, 1966.

Carlson, T.R. The relationship between speed and accuracy of comprehension. Journal of Educational Research, 1949, 42, 500-512.

Chall, J.S. Readability: an appraisal of research and application. Bureau of Educational Research Monographs, Ohio State, 1958, No. 34.

Chandler, T.A. Reading disability and socio-economic status. Journal of Reading, 1966, 10, 5-21.

Chomsky, C. The acquisition of syntax in children from 5 to 10. Unpublished doctoral dissertation, Harvard University, 1968.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: M.I.T. Press, 1965.

Chomsky, N. Syntactic structures. The Hague: Mouton & Co., 1957.

Clark, H.H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404.

Cleland, D.L. A construct of comprehension. In J.A. Figurel (Ed.), Reading and inquiry. International Reading Association: Conference Proceedings, Vol. X: Newark, Delaware. 1965, 59-64.

Cleland, D.L. Needed improvement in research design in reading. In J.A. Figurel (Ed.) Improvement of reading through classroom practice. International Reading Association Proceedings, Vol.9. Newark, Delaware; 1964, 244-249.

Clymer, T. What is reading?: some current concepts. In Helen M. Robinson (Ed.), Innovation and change in reading instruction. 67th Yearbook, National Society for Study in Education, Part II: Chicago: University of Chicago Press: 1968, 7-30.

Cooper, Bernice. An analysis of reading achievement of white and negro pupils in certain public schools of Georgia. School Review, 1964, 72, 467-471.

Cronbach, L.J. & Meehl, P.E. Construct validity in psychological tests. Psychology Bulletin. 1955, 52, 281-302.

Davidson, R.E. & Dollinger, L.E. Syntactic facilitation of paired associate learning: Deep structure variations. Journal of Educational Psychology, 1969, 60, 434-438.

Davis, F.B. Fundamental factors of comprehension in reading. Psychometrica, 1944, 9, 185-197.

Davis, F.B. Measurement of improvement in reading skill courses. In E. Bliesmer (Ed.) Problems, programs, and projects in college-adult reading. Eleventh yearbook of the National Reading Conference, Milwaukee: 1962, 30-40.

Davis, F.B. Research in comprehension in reading. Reading Research Quarterly, 1968, 3, 499-545.

Duker, S. Listening and reading. Elementary School Journal, 1965, 65, 321-329.

Durrell, D.D., Improving reading instruction. New York: World Book Co. 1956.

Fodor, J. Psychological explanation: An introduction to the philosophy of psychology. New York: Random House, 1968.

Fodor, J.A., & Garrett, M. Some reflections on competence and performance. in J. Lyons & R.J. Wales (Eds.) Psycholinguistic papers. Edinburgh: Edinburgh University Press, 1966, 135-179.

Fodor, J.A., Garrett, M. & Bever, T.G. Some syntactic determinants of sentential complexity, II: Verb structure. Perception and Psychophysics, 1968, 3, 453-461.

Gans, Roma. A study of critical reading comprehension in the intermediate grades. Teachers College Contributions to Education, 1940, No. 811.

Gates, A.I. Sex differences in reading ability. Elementary School Journal, 1961, 61, 431-434.

Guilford, J.P. Psychometric methods. New York: McGraw-Hill, 1951.

Hall, W. & Robinson, F. An analytical approach to the study of reading skills. Journal of Educational Psychology, 1945, 36, 429-442.

Harootunian, B. Intellectual abilities and reading achievement. Elementary School Journal, 1966, 66, 386-392.

Holmes, J.A. & Singer, H. Speed and power of reading in high school. Cooperative Research Monograph #14, U.S.O.E. 1966.

Huey, E.B. The psychology and pedagogy of reading. Macmillan: 1908 (Republished: Cambridge: M.I.T. Press, 1968).

Jacobs, R. & Rosenbaum, P. English transformational grammar. Waltham, Mass: Blaisdell. 1968.

Jacobson & Johnson. The relationship of attitudes to reading comprehension in the intermediate grades. Paper presented at the American Educational Research Association meeting, 1967.

Jenkinson, Marion. Selected process and difficulties of reading comprehension. Unpublished doctoral dissertation. University of Chicago, 1957.

Kerfoot, J.K. Comprehending comprehension. In J.A. Figurel, (Ed.) Improvement of reading through classroom practice, International Reading Association Proceedings, Vol. 9. Newark, Delaware: 1964, 74-76.

Kingston, A. A conceptual model of reading comprehension. In Bleismer & Kingston (Eds.), Phases of college and adult reading, 10 yearbook: National Reading Conference, 1961.

Klare, G. The measurement of readability. Ames, Iowa: Iowa State University Press, 1963.

Kuhn, T.S. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.

Langacker. R.W. Language and its structure. New York: Harcourt, Brace & World, 1967.

Langsman, R.S. A factorial analysis of reading ability. Journal of Experimental Education, 1941, 10, 51-63.

Levelt, W.J.M. Psychological representations of syntactic structures. In Bever T.G. & Weksel W. (Eds.) Studies in psycholinquistics, in press.

Levin, H. & Wanat, S. The eye-voice-span: Reading efficiency and syntactic predictability. Studies in oral reading: XI Unpublished manuscript. Cornell, University, 1967.

Loban, W. Language ability: grades 7, 8, 9. Cooperative Research Monograph, No. 18, U.S.O.E., 1966.

McCall, W.A. & Crabbs, L. Standard test lessons in reading. New York: Teachers College, Columbia University Press, 1961.

Mehler, J., & Carey, P. Role of surface and base structure in the perception of sentences. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 335-338.

Miller, G.R. & Coleman, E. A set of thirty six prose passages calibrated for complexity. Journal of Verbal Learning and Verbal Behavior, 6, 1967, 851-854.

Neville, D.,Pfost, P. & Doobs, V. The relationship between anxiety and silent reading gain. American Educational Research Journal, 4, Jan. 1967, 45-51.

Olson, A.V. School achievement, reading ability, and specific visual perception skills in the third grade. Reading Teacher, 19, 1966, 490-493.

Olson, W.C. Reading as a function of the total growth of the child. In W.S. Gray (Ed.), Reading and pupil development, Chicago: University of Chicago Press, 1940, 223-237.

Piekarz, J. Individual differences in interpretive responses in reading. Unpublished doctoral dissertation. University of Chicago, 1954.

Potter, Thomas C. A taxonomy of cloze research, part 1: Readability and reading comprehension. Southwest Regional Laboratory for Educational Research and Development, Inglewood, California, 1968.

Schell, L.M. Distinguishing fact from opinion. Journal of Reading, XI, Oct. 1967, 5-9.

Schutz, R.E. Testing theories of reading instruction. In John Bormuth (Ed.), Highlights, the 1966 Institutes. International Reading Association, Newark, Delaware: 1966.

Smith, D.E.P. Reading comprehension: a proposed model. In O. Causey (Ed.), Research and evaluation in college reading. Ninth Yearbook: National Reading Conference. Fort Worth: Texas Christian University Press, 1960, 21-27.

Spache, G. Toward better reading. Champaign, Illinois: Garrard, 1967.

Sparks, J. & Mitzel, H. A reaction to Holme's basic assumption underlying the substrata-factor theory. Reading Research Quarterly, 1 1966, 137-146.

Tatham, S.M. The effect of material written with certain oral language patterns on reading comprehension. Paper presented at the American Educational Research Association meeting, 1967.

Taylor, W.L. Cloze procedure: a new tool for measuring readability. Journalism Quarterly, 30, 1953, 414-438.

Thorndike, E.L. Reading as reasoning: a study of mistakes in paragraph reading. Journal of Educational Psychology, VIII, 1917, 323-332.

Thurstone, L.L. Note on a reanalysis of Davis' reading tests. Psychometrica, 11, 1946, 185-188.

Tinker, M.A. Speed versus comprehension in reading by level of difficulty. Journal of Educational Psychology, 30, 1939, 81-94.

Tremont, J. Cloze procedure as a method for measuring readability of text, reading comprehension of groups, and idea density in instructional materials. Unpublished qualifying paper, Harvard Graduate School of Education, 1967.

Vehar, M.A. Extro-version, introversion, and reading ability. Reading Teacher, 21, 1962, 357-360.

Vernon, P. The determinants of reading comprehension. Educational and Psychological Measurement, 22, 1962, 269-286.

Wanner, R. On remembering, forgetting, and understanding sentences: a study of the deep structure hypotheses. Unpublished doctoral dissertation, Harvard University, 1968.

Wardhaugh, R. Reading: a linguistic perspective. New York: Harcourt, Brace & World, 1969.

Weaver, W. & Kingston, A. Factor analysis of the cloze procedure and other measures of language ability. Journal of Communication, 13, 252-261.

Wolf, W., King, M. & Huck, C. Teaching critical reading to elementary school children. Reading Research Quarterly, 3, 1968, 435-498.

VITA

1955-1959	Brandeis University	B.A. June 1959
1959-1962	Massachusetts College of Optometry	B.S. June 1962 O.D. 1964
1962-1963	School of Education Boston University	Ed.M. June 1963
1963-1964	Teacher, Reading, Pollard Junior High School Needham, Massachusetts	
1964-1965	Educational Program Specialist Action for Boston Community Development, Boston, Massachusetts	
1965-1966	Graduate School of Education Harvard University	Ed.M. June 1966
1966-1969	Research Assistant Graduate School of Education Harvard University	
1969-1970	Associate in Education Graduate School of Education Harvard University	